



OULUN YLIOPISTO
UNIVERSITY of OULU

OULU BUSINESS SCHOOL

Lauri Heikka

**MOMENTUM: AUTOCOVARIANCES, CROSS-COVARIANCES OR UNCONDITIONAL
EXPECTED RETURNS?**

Master's Thesis

Finance

December 2019

Unit Department of Finance			
Author Lauri Heikka		Supervisor Hamed Salehi	
Title Momentum: autocovariances, cross-covariances or unconditional expected returns?			
Subject Finance	Type of the degree Master	Time of publication December 2019	Number of pages 104
<p>Abstract</p> <p>Returns to momentum strategies can be decomposed into three sources of return: positive autocovariances in returns, negative cross-covariances in returns and cross-sectional variation in unconditional expected returns across assets. While theoretical literature on momentum generally assumes that positive autocovariances drive momentum returns, empirical literature presents inconsistent evidence on the importance of each component in explaining momentum returns. However, this prior literature ignores a key empirical issue related to short return histories and extreme return observations. These extreme returns cause a negative bias to sample estimates of return autocovariances and cross-covariances, and a positive bias to the cross-sectional variation in unconditional expected returns. Furthermore, prior literature focuses on portfolios of stocks instead of individual stocks, because the decomposition requires estimating a cross-covariance matrix, which is difficult for individual stocks with short and non-overlapping return histories.</p> <p>I propose a novel, strategy-based decomposition, which allows for estimating the contribution of each component to momentum without bias in the presence of extreme returns and non-overlapping return histories. Empirical evidence from the strategy-based decomposition in a sample of US individual stocks and portfolios of US stocks extending from 1926 to 2018 suggests, that positive return autocovariances are the most important driver of momentum. This evidence is consistent with most behavioral theories on momentum but does not preclude a rational interpretation. When the contributions are allowed to vary over time, positive autocovariances consistently remain the most important driver of momentum returns throughout the sample period. However, the time-series predictability in individual stock returns slowly erodes after momentum is initially documented in the literature in 1993, potentially explaining the poor recent performance of momentum.</p>			
Keywords Asset pricing – Anomalies – Momentum			
Additional information			

CONTENTS

1	INTRODUCTION.....	6
2	LITERATURE REVIEW	12
2.1	Efficient markets and the Capital Asset Pricing Model	12
2.2	Momentum.....	13
2.3	Related phenomena	15
2.4	Explanations for the momentum premium.....	17
2.4.1	Behavioral theories	17
2.4.2	Rational theories	19
2.4.3	Factor explanations	20
2.5	Determinants of momentum returns	22
2.5.1	Cross-sectional determinants of momentum.....	23
2.5.2	Momentum over time and the market cycle	26
2.5.3	Momentum and limits to arbitrage.....	27
2.6	Momentum and return autocovariance	28
2.6.1	The decomposition and theories on momentum	29
2.6.2	Empirical evidence on the decomposition of momentum.....	32
3	RESEARCH QUESTIONS AND CONTRIBUTION TO THE LITERATURE	38
4	DATA & METHODS	42
4.1	Data.....	42
4.2	Relative-weight momentum and the Lo and MacKinlay (1990) decomposition	44
4.3	Decomposition estimation.....	47
4.4	Biased estimation and a bootstrap experiment	49
4.4.1	Adjusting for extreme return observations with winsorization.....	52
4.5	Strategy-based decomposition.....	53

4.5.1	Estimation methods for the strategy-based decomposition	57
5	EMPIRICAL RESULTS	61
5.1	Momentum returns and the unconditional decomposition	61
5.2	Bootstrap evidence of biased estimation	64
5.3	Decomposition with winsorized returns	67
5.4	Strategy-based decomposition of momentum returns	69
5.4.1	Strategy returns using full sample estimates	70
5.4.2	Strategy returns using out-of-sample estimates	71
5.4.3	Return regressions	73
5.5	Time-variance in the return mechanism	76
5.6	Robustness checks	81
5.6.1	Decomposition across market capitalizations	81
5.6.2	Robustness to scaling the strategy weights	83
5.6.3	Robustness to methods of estimating the unconditional expected return	86
5.6.4	Other formation and holding periods of momentum	88
6	SUMMARY	91
	REFERENCES	94
	APPENDICES	
	Appendix 1 Strategy and decomposition derivations	100

FIGURES

Figure 1.	10-year rolling annualized percentage returns of component strategies	78
Figure 2.	Cumulative returns of momentum and component strategies.	78

TABLES

Table 1. Summary statistics of the individual stock and portfolio data samples from 1926 to 2018.....	43
Table 2. Momentum strategy returns and unconditional decomposition results, in the sample period from 1926 to 2018, all datasets	62
Table 3. Average momentum strategy returns and decomposition estimates from 500 bootstrap replications under the null hypothesis of no time-series dependence in returns, 1926 to 2018 sample period	65
Table 4. Momentum returns and decomposition estimates in winsorized returns 1926-2018	68
Table 5. Returns to momentum and component strategy returns, using full-sample estimates of the unconditional expected return from the 1926 to 2018 sample period	70
Table 6. Returns to momentum and component strategy returns, using expanding-window estimates of the unconditional expected, 1926 to 2018 sample period	72
Table 7. Regression coefficients and intercepts for return regressions	74
Table 8. Decomposition of momentum returns in different sample periods, momentum in individual stocks	80
Table 9. Decomposition of momentum across market capitalizations	82
Table 10. Returns to component strategies using different methods of scaling the investment weights in the strategies, from 1926 to 2018	84
Table 11. Momentum and component strategy returns in individual stocks from 1926 to 2018, using different estimates of the unconditional expected return to form the strategy weights	87
Table 12. Momentum and component strategy returns in individual stocks from 1926 to 2018 with different holding periods and formation periods used to form the strategy weights.....	89

1 INTRODUCTION

Momentum, the tendency of assets that have performed well recently to outperform assets that have done poorly recently, is a robust and pervasive phenomenon. Jegadeesh and Titman (1993) find, that strategies buying stocks with high returns over the past 3 to 12 months and selling stocks with low returns over the same horizon, produce statistically significant and economically large positive returns over subsequent months. Subsequent literature documents momentum being a robust and pervasive phenomenon that is present in both markets outside the US, as well as asset classes outside of stocks.¹ The robustness of the evidence on momentum has led to an extensive amount of theories proposed in the literature aspiring to explain momentum. These theories can be roughly divided into behavioral and rational theories. Behavioral theories on momentum posit, that momentum arises due to irrational behavior by market participants.² A common interpretation is that momentum is caused by underreaction to firm- or asset-specific information by market participants, which in some theories, is followed by an overreaction. Theories based on rational expectations interpret momentum as compensation for some underlying risks, not accounted for by traditional asset pricing models.³

¹ Moskowitz and Grinblatt (1999) find momentum in industry portfolios, and Lewellen (2002) documents similar evidence with portfolios sorted on market capitalization and book-to-market ratios. Momentum is also present in markets outside the US (Rouwenhorst (1998), Asness, Moskowitz & Pedersen (2013)), as well as in asset classes other than stocks (Asness, Moskowitz & Pedersen (2013), Jostova, Nikolova, Philipov & Stahel (2013)). Moskowitz, Ooi and Pedersen (2012) document a related time-series momentum phenomenon and find it to be persistent across multiple asset classes and markets.

² Daniel, Hirshleifer and Subrahmanyam (1998), Luo, Subrahmanyam and Titman (2018) and Barberis, Schleifer and Vishny (1998) attribute momentum to different cognitive biases, which cause investors to underweight new publically available information. Hong and Stein (1999) suggest a model with both under- and overreaction, with gradual diffusion of information among investors. Grinblatt and Han (2005) and Frazzini (2006) argue that the disposition effect causes excess selling pressure on past winners and undersupply on past losers, which leads past winners (losers) to become under- (over-) valued.

³ Berk Green and Naik (1999) and Chordia and Shivakumar (2002) suggest that systematic exposures to a set of macroeconomic risk factors explain momentum. Johnson (2002), Sagi and Seasholes (2007) and Liu and Zhang (2008) relate momentum to time-varying risk exposures, positing that a firms cash flows become more risky after facing recent growth. Ahn Conrad and Dittmar (2003) provide another explanation based on time-varying risk exposures, while Holden and Subrahmanyam (2002) build a rational model of momentum based on information asymmetry. Another branch of the literature proposes new risk factors that might explain momentum at least to a degree (see e.g. Hou, Xue & Zhang (2015), Pastor and Stambaugh (2003), Ruenzi and Weigert (2018) and Sadka (2006)).

Recently, Subrahmanyam (2018) calls for more empirical work towards thinning out the large number of competing theories on momentum. Testing many theories jointly is difficult, however, because each theory generates a wide range of testable predictions. This thesis focuses on one very specific aspect of momentum, which is relevant to all theories: the return mechanism that causes momentum. Lo and MacKinlay (1990) suggest an analytical decomposition of strategies that weight assets in proportion to their past returns. This decomposition reveals that returns to momentum can arise in three ways: 1. Returns have positive autocovariances i.e. an asset's past returns positively predict its subsequent returns. 2. Returns have negative cross-covariances, such that high (low) returns to some asset predicts low (high) returns to some other asset. 3. Momentum strategies on average weight positively on assets with higher unconditional expected returns and negatively on assets with lower unconditional expected returns. If the cross-sectional variation in unconditional expected returns is large, momentum generates significant positive returns.

Virtually all behavioral theories in the literature, but some rational ones as well, imply that momentum is driven by positive autocovariances in returns. Some rational theories suggest, that variation in unconditional expected returns drives momentum instead. Negative cross-covariances have largely been ignored by the literature as a potential driver of momentum. However, prior empirical evidence on the return decomposition is inconsistent with positive autocovariances driving momentum. Conrad and Kaul (1998), Bulkley and Nawosah (2009) and Park and Kim (2014) find evidence in support of cross-sectional variation in unconditional expected returns driving momentum in individual stocks. Lewellen (2002) studies portfolio-level momentum and finds evidence consistent with momentum being driven by negative cross-covariances. Du (2012) finds similar evidence for momentum in individual stocks. Pan, Liano & Huang (2004) find positive autocovariances in industry portfolios only for lags up to three months but negative autocovariances at longer horizons.

Prior literature on decomposing momentum returns focuses mainly on momentum in portfolios of stocks, because the standard decomposition requires an estimate of the cross-covariance matrix of returns. Estimating the cross-covariance matrix for individual stocks is difficult, because there are a large number of individual stocks with short and non-overlapping return histories. I contribute to prior literature by

proposing two methods to tackle this problem. The first one involves estimating the decomposition by limiting the sample of stocks into stocks with 24 months of common return history and estimating the decomposition in a moving window. The second method is a strategy-based decomposition of momentum returns, novel to the literature. I analytically solve for the weights to three investment strategies that target each of the three return mechanisms related to momentum separately. Unlike the standard decomposition, the strategy-based decomposition does not require an estimate of the cross-covariance matrix and can be estimated with relative computational ease.

The inconsistency between empirical evidence and theories on the return mechanism is puzzling. Chen and Hong (2002) argue that prior findings of negative autocovariances are caused by investors overreacting to market-wide news, while underreacting to firm-specific information. However, they provide only minimal empirical support. I provide a different interpretation of prior findings. Prior literature uses sample estimates of autocovariances, cross-covariances and the cross-sectional variation in unconditional expected returns. These sample estimates have been shown to be sensitive to extreme observations, potentially causing bias in estimating the decomposition.⁴

I empirically test for the presence of biases in the decomposition in US individual stocks and portfolios of US stocks, by conducting a bootstrap experiment. I sample individual return observations to generate return series' with no time-series predictability in returns and form momentum strategies on these scrambled returns. The returns on these strategies should only depend on the cross-sectional variation in unconditional expected returns. However, decomposing the returns to these strategies reveals a significant negative contribution by return autocovariances and significant positive contribution by return cross-covariances. This evidence suggests that sample estimates of the contribution of autocovariances are negatively biased, while estimates

⁴ Jegadeesh and Titman (2002) show that sample estimates of the cross-sectional variation in unconditional expected returns are positively biased by extreme return observations and short return histories. Chan (1992, 1995) and Deutsch, Richards and Swain (1990) show that very few outlier observations can cause considerable bias in sample estimates of autocovariances.

of the contributions of cross-covariances are positively biased. In addition, I replicate the results of Jegadeesh and Titman (2002), who show that sample estimates for the cross-sectional variation in unconditional expected returns are positively biased. When returns are sampled without replacement, such that extreme return observations cannot enter into both the holding and formation periods, the returns of the scrambled momentum strategies are very small and statistically insignificant for both individual stocks and industry portfolios, suggesting that the variation in unconditional expected returns does not drive momentum returns.

Re-estimating the decomposition and the bootstrap experiment using winsorized returns, where extreme returns are pulled in, provides moderate evidence suggesting that these biases are caused by extreme return observations. However, the strategy-based decomposition provides a more intuitive way to control for the biases. The strategy-based decomposition allows for an interpretation of the biases in the decomposition as forward-looking biases. Using ex-ante available information to form these component strategies results in unbiased estimates of the decomposition, regardless of whether the biases are caused by extreme returns or some other phenomena.

I empirically examine the strategy-based decomposition in a sample of US individual stocks and portfolios of US stocks, sorted by industry, size and book-to-market ratios extending from 1928 to 2018. My main findings support the view that momentum is mostly driven by positive autocovariances in returns. Using the strategy-based decomposition, positive return autocovariances are the only driver of momentum that an investor can target to earn significant positive returns using information available ex-ante. Furthermore, the cross-covariance strategy is negatively correlated with both momentum and the autocovariance strategy, while the autocovariance strategy is positively correlated with momentum. These correlations are strong and provide further evidence that negative cross-covariances do not drive momentum returns. However, when momentum returns are regressed against component strategy returns, momentum has a large positive regression intercept over all component strategies, suggesting that these component strategies do not capture momentum in its entirety.

Momentum driven by positive autocovariances is consistent with most behavioral theories on momentum. Furthermore, a number of rational theories predict this relationship as well. Thus, while my results address some puzzling empirical results in prior literature, they do not significantly reduce the number of plausible theories for momentum. Because different theories on momentum generate varying testable hypotheses, evaluating theories jointly remains an important challenge in understanding momentum.

My main results are robust across market capitalizations of individual stocks and a variety of other methodological choices. Furthermore, results for industry portfolios and individual stocks are consistent with each other, suggesting that momentum in individual stocks and industry momentum are similar phenomena. Evidence for size and book-to-market sorted portfolios suggests that momentum in these characteristic-sorted portfolios is distinct from momentum in individual stocks and industry portfolios and more driven by variation in unconditional expected returns, related to size and value factor returns. A further examination of momentum in these characteristic sorted portfolios is left for future research, however.

Another advantage of the strategy-based decomposition is that it naturally allows for time-variance in the return mechanism. Exploring time-variance, I find that positive autocovariances, are consistently the most important driver of momentum returns in individual stocks throughout the sample period. While negative return cross-covariances seem to generate positive momentum returns at times in the latter part of the sample, negative return autocovariances dominate the cross-covariances resulting in poor performance of momentum strategies. I also uncover some evidence suggesting that time-series predictability in returns steadily erodes after 1990, placing concern over the profitability of momentum strategies in the future. My findings are consistent with McLean and Pontiff (2016), who find, that cross-sectional return predictability tends to attenuate post-publication.

The strategy-based decomposition also reveals a novel reversal phenomenon in individual stocks akin to the long-term reversal effect first documented by DeBondt and Thaler (1985). I find that a contrarian strategy that invests long in stocks with low lifetime historical returns and short in stocks with high lifetime historical returns earns

statistically significant and economically large positive returns consistently throughout the sample period. The returns to this contrarian strategy are similar when estimates of unconditional expected returns are formed using full lifetime returns, excluding the holding period return, and when estimates are formed using only post-holding period returns. This suggest that stocks with extremely high (low) individual monthly return realizations earn lower (higher) than average returns over their lifetime. My findings add to prior empirical evidence on investors requiring a return premium for holding stocks with negatively skewed returns over stocks with positively skewed returns (Bali, Cakici & Whitelaw, 2011 and Boyer, Mitton & Vorkink, 2010).

The remainder of the thesis proceeds as follows. I review prior literature on momentum, including the empirical evidence, the most important theories and empirical tests of these theories in section 2. In section 3, I outline my research questions and my contribution to prior literature. I describe my data and main methods in section 4. In section 5, I present my empirical results. I conclude in Section 6.

2 LITERATURE REVIEW

In this section, I review prior literature related to cross-sectional momentum in stock returns, henceforth referred to as momentum. I begin with a brief overview of early work on cross-sectional asset-pricing and the evidence on the momentum anomaly. I then discuss different explanations for the momentum phenomenon suggested in the literature and provide a summary of literature examining different cross-sectional and time-series attributes and determinants of momentum strategies. I conclude the literature review with a discussion of the decomposition of momentum returns into distinct drivers of returns and a review of prior evidence related to the decomposition.

2.1 Efficient markets and the Capital Asset Pricing Model

Much of the early paradigm of asset pricing literature builds on the efficient market hypothesis. In his influential survey of efficient markets literature, Fama (1970) defines efficient market as one where prices fully reflect all available information. When all available information is reflected in prices immediately, prices follow an unpredictable “random walk”, and no systematic investment strategy can earn abnormal returns. Under efficient markets, the return available to any investor is always proportionate to the risk taken on by the investor.

A large body of empirical work is dedicated to examining the efficient market hypothesis. Testing market efficiency has proven difficult, however. Testing market efficiency requires a benchmark model that expected returns should follow if markets are efficient. Because of this, any test of market efficiency faces a joint hypothesis problem. As Fama (1970) points out, tests of market efficiency simultaneously test market efficiency and the validity of the benchmark model of asset price formation. Thus, it is difficult to assess whether a finding indicates inefficiencies in the market, or whether the benchmark model is an insufficient model of the risk-reward relationship in expected returns.

In an early effort to quantify the risk-reward relation and to provide a benchmark for tests of market efficiency, Sharpe (1964) and Lintner (1965) develop the Capital Asset Pricing Model (from now on CAPM for short). The central assertion of the CAPM is

that the cross-sectional variation in expected returns can be explained by a function of each assets' covariance with the aggregate market return. In theory, the market return should include returns on all investable assets, including not just stocks and bonds, but other investments such as real estate as well. However, most empirical research on stock returns typically use a value-weighted portfolio of stocks traded in the U.S. as a proxy for this market factor. The CAPM has been a widely used benchmark equilibrium model of asset prices in subsequent literature.

Despite being an important landmark in asset pricing theory, empirically the CAPM has not been a success. For example, Fama and French (1992, 1993) do not detect the connection between market betas and future stock returns predicted by the CAPM. Subsequent literature also identifies predictability in the cross-section of stock returns not related to market betas. These patterns are often referred to as anomalies. Harvey, Liu and Zhu (2016) explore asset pricing literature and find over 300 variables suggested to explain the cross-section of stock returns. Some of the earliest and most well-known studies show that market capitalization (Banz, 1981), book-to-market-ratios (Rosenberg, Reid, & Lanstein, 1985), and past returns (Jegadeesh and Titman, 1993) all predict variation in future returns. In light of these findings, many improvements upon the CAPM have been suggested in the literature. For example, Fama and French (1992, 1993) add two additional factors to the CAPM, based on market capitalization and book-to-market ratios. The resulting three-factor model is widely used as a benchmark model in subsequent anomalies literature.

2.2 Momentum

Momentum is one of the most studied phenomena in the anomalies literature. In the context of stock markets, momentum can be broadly defined as the tendency of stocks with higher returns over the intermediate horizon (3 to 12 months), to outperform stocks with lower returns over the same horizon in subsequent months for up to a year. The discovery of momentum in stock returns is most often credited to Jegadeesh and Titman (1993), although similar concepts were explored earlier, for example by Levy (1967) with his relative strength strategies. In their seminal paper, Jegadeesh and Titman (1993) consider strategies that sort stocks into portfolios based on their realized return over the past 3 to 12 months and then buy the portfolio with the highest recent

realized returns and sell short the portfolio with the lowest returns. They find that the resulting zero-investment portfolios yield economically large and statistically significant returns that are not explained by the CAPM. Jegadeesh and Titman (1993) study various formation periods, as well as various holding periods. Their strategies also hold multiple portfolios simultaneously. The largest returns to an individual strategy documented by Jegadeesh and Titman (1993) are 1.31% per month, with a 12-month formation period and a holding period of 3 months following immediately after the formation period. The approach based on overlapping portfolios in Jegadeesh and Titman (1993) is uncommon in subsequent literature. Asness, Moskowitz and Pedersen (2013) posit that the most standard approach in the literature is to form momentum portfolios based on 11-month cumulative return of a stock, during the period beginning 12 months prior and ending one month prior to portfolio formation. The rationale for skipping one month between the holding period and the momentum period is the tendency of previous performance over periods from one week to one month to be negatively related to following performance. This short-term reversal effect is documented by Jegadeesh (1990) and Lehmann (1990). DeBondt and Thaler (1985) uncover another similar reversal effect, when predicting future returns with longer-term (one to five years) past performance. Bali, Engle, and Murray (2016 p. 209-219) show that the choice of formation period and holding period has some effect on the magnitude of momentum returns. The 11-month cumulative return with one month skipped in between the formation and holding periods generates the highest returns in their sample of US stocks from 1963 to 2012.

Following Jegadeesh and Titman (1993), a broad literature posts evidence of momentum being a persistent and pervasive phenomenon. Rouwenhorst (1998) replicates Jegadeesh and Titman (1993) in 12 European countries and finds similar results. Griffin, Ji and Martin (2003) and Chui, Titman and Wei (2010) look for evidence around the world and find significant and positive momentum returns in most countries. Both find, that momentum tends to be weaker in Asian countries. Moskowitz and Grinblatt (1999) and Lewellen (2002) document momentum in portfolios of U.S. stocks. Asness, Moskowitz and Pedersen (2013) show that momentum exists not just across international stock markets, but in and across different asset classes, such as equity indices, currencies, commodity futures and government bonds as well. Jostova, Nikolova and Philipov (2013) document

momentum in non-investment grade corporate bonds. Jegadeesh and Titman (2001) and Israel and Moskowitz (2013) find that momentum has remained profitable in post-publication periods as well.

Momentum returns tend to also survive most common risk-adjustments. Jegadeesh and Titman (1993) show, that momentum returns are not explained by exposure to market risk. Fama and French (1996, 2008, 2016) find that momentum profits remain strong and pervasive even, when adjusting for risk-exposures based on their three-factor and five-factor models. Lewellen and Nagel (2006) and Bali and Engle (2010) document failures of conditional and intertemporal versions of the CAPM to explain momentum. Carhart (1997) augments the Fama and French (1993) three-factor model with a momentum factor. According to Bali, Engle and Murray (2016, p. 239) the resulting four-factor model is the most widely used benchmark risk model in empirical asset pricing literature.

2.3 Related phenomena

Moskowitz and Grinblatt (1999) document a significant momentum effect in industry portfolio returns. They empirically link this phenomenon to stock-level momentum, by comparing industry momentum strategies, and strategies, where ‘industry’ is assigned to stocks at random. They find, that industry momentum strategies produce significant positive returns, while the random industry strategies do not, concluding that momentum in individual stocks is primarily driven by momentum in industries. Lewellen (2002) documents a similar momentum effect in portfolios sorted on market capitalization and book-to-market ratios. Lewellen (2002) finds that cross-sectional momentum between these characteristic-sorted portfolios remains significantly positive for up to 17 months after formation. This is in contrast with industry and individual stock momentum returns, which in Lewellen’s (2002) results tend to turn negative after 9 to 11 months. Lewellen (2002) further examines, whether momentum in these portfolios is present in risk-adjusted returns using the Fama and French (1993) three-factor model. He finds that returns to the factors in the three-factor model of Fama and French exhibit significant momentum. Furthermore, momentum in these factors seems to account for most of the momentum that the Fama and French portfolios exhibit. Industry portfolios on the other hand, generate significant and

persistent momentum even when strategies are formed using the residuals of the Fama and French three-factor model. Blitz, Huij & Martens (2011) investigate this relationship of the Fama and French factors using individual stocks. They report, that momentum strategies constructed based on idiosyncratic, stock-specific returns dominate ones constructed on total returns or ones based on factor exposures.

Grundy and Martin (2001) re-examine the claims in Moskowitz and Grinblatt (1999) and show that industry momentum is to a large extent dependent on returns in the most recent month. When considering industry momentum strategies that skip the most recent month, they find that the outperformance of true industry momentum to momentum with random industries is no longer significant. Based on these findings, they posit that stock-level momentum cannot entirely explain momentum in individual stocks. Chordia and Shivakumar (2002) link both stock-level momentum and industry momentum to common macroeconomic factors. They show that the way industry momentum and individual stock momentum are related to these macroeconomic variables are independent of each other. Chordia and Shivakumar (2002) posit that this finding is supportive of stock level and industry momentum being separate phenomena. However Griffin and Martin (2003) find no supporting evidence for the empirical findings Chordia and Shivakumar (2002) when looking at international markets.

An extensive literature documents earnings momentum, predictability in returns based on past earnings. Research examining this relation include Latane and Jones (1979), Bernard and Thomas (1989) and Chan, Jegadeesh and Lakonishok (1996). The key finding in this literature is, that firms with high earnings surprises tend to outperform those with negative earnings surprises over subsequent periods. Studying the link between momentum and earnings momentum, Chan et al. (1996) find that earnings momentum and momentum are positively correlated but also contain complementing idiosyncratic information on subsequent performance. They posit that neither phenomenon is subsumed by the other. Contrastingly, Chordia and Shivakumar (2006) find that a portfolio long on highest and short on lowest earnings surprises captures most of the returns to a momentum strategy.

Moskowitz, Ooi and Pedersen (2012) document time-series momentum. Time-series momentum differs from cross-sectional momentum in that time-series momentum focuses purely on an assets own recent return, instead of its return relative to other assets. Moskowitz et al. (2012) find time-series momentum to be pervasive across markets and asset classes and show that a time-series momentum factor captures completely the returns to traditional cross-sectional momentum. Bird, Gao and Yeung (2016) find evidence across 24 markets that time-series momentum is stronger than cross-sectional momentum. Goyal and Jegadeesh (2016) find that time-series momentum is present in individual US stocks as well. However, Goyal and Jegadeesh (2018) argue that results for time-series momentum strategies cannot be interpreted like cross-sectional momentum strategies. This is because, unlike cross-sectional momentum, time-series momentum strategies do not construct market-neutral long-short portfolios. Instead time-series momentum takes on a time-varying net long position in risky assets. Goyal and Jegadeesh (2018) show that controlling for this time-varying long position markedly reduces returns to time-series momentum.

2.4 Explanations for the momentum premium

The robust evidence on momentum has generated an extensive literature of different theories aspiring to explain the effect. These theories can be divided into two competing views. Some propose rational theories, positing that momentum is compensation for risks that standard benchmark models do not account for. Others attribute momentum to cognitive biases and heuristics that affect investor behavior.

2.4.1 Behavioral theories

In behavioral explanations, momentum most often arises from two sources: underreaction and a delayed overreaction. These effects are not mutually exclusive and can coexist and reinforce each other to create momentum. Daniel, Hirshleifer and Subrahmanyam (1998) propose a model, in which momentum arises through overconfidence and self-attribution bias. They model a representative investor who is overconfident in their own ability and attributes their successes to skill and failures to luck. These biases cause investors overreact to private information that they themselves have gathered, while underreacting to public information. Daniel et al.

(1998) show that momentum in their theory can arise from both underreaction to public information and continued overreaction to private information through positive autocovariances in returns. As the overreaction reverts, autocovariances turn negative, consistent with long-term reversals. More recently, Luo, Subrahmanyam & Titman (2018) provide another theory based on overconfidence. While Daniel et al. (1998) focus on a representative investor, Luo et al. (2018) model heterogeneous investors. They note that their model improves upon Daniel et al. (1998) in that it predicts post-earnings announcement drift as well.

Barberis, Shleifer and Vishny (1998) provide another model based on underreaction by a representative investor. They suggest that conservatism, the tendency of an agent to insufficiently revise their beliefs based on new information, causes underreaction. Overreaction in turn is caused by the representativeness heuristic, which causes investors to mistakenly extrapolate persistent past news as representative of an underlying pattern. In their model, investors transition between two regimes, in which they either underreact or overreact to new information. The underreaction regime causes momentum and subsequent overreaction then exacerbates this effect, giving rise to reversals in the long-term.

Hong and Stein (1999) instead model a universe with two types of investors, newswatchers and momentum traders. Momentum in their model arises from the interaction between these two groups of investors. Private information is gradually incorporated among the population of newswatchers, who are indifferent to past returns, generating underreaction to new information. The momentum traders instead do not incorporate new information into their trading, but instead condition purely on past returns. The initial underreaction makes it profitable for the momentum traders to trade based on past returns. Hong and Stein (1999) show, that this not only corrects the underreaction, but is bound to cause an eventual overreaction, as momentum traders continue to trade on past returns even after news is fully incorporated into prices.

Antoniou, Doukas and Subrahmanyam (2013) extend on the model in Hong and Stein (1999) establishing a connection between investor sentiment and underreaction. They hypothesize that, when new information arrives that contradicts their prior sentiment,

the newswatchers experience cognitive dissonance resulting in a more pronounced underreaction. Cognitive dissonance occurs when an agent experiences discomfort caused by receiving information that contradicts their personal beliefs. Hong, Stein & Yu (2007) propose a model, where investors use oversimplified models in forming their opinions of stocks, ignoring some available information. This causes them to make persistent forecast errors, while ignoring true information, leading to an underreaction effect.

Another potential source for underreaction that might cause momentum over the intermediate term is the disposition effect, the tendency of investors to hold on to assets that have lost value, while being eager to sell assets that have increased in value. Grinblatt and Han (2005) argue that the disposition effect causes excess selling pressure in stocks that have experienced positive returns recently compared to stocks that have experienced losses. This occurs, because investors seek to realize capital gains in the winner but want to hold on to the loser, causing past winners to become undervalued and past losers to become overvalued. Momentum then arises as prices gradually correct towards fundamentals. Grinblatt and Han (2005) validate their model empirically, by showing that a proxy of the marginal investors' unrealized capital gains, constructed using return and trading volume data, predicts future returns better than the past return of a stock. Frazzini (2006) finds similar results using a proxy of unrealized capital gains, constructed from mutual fund holdings data.

2.4.2 Rational theories

Theories based on rational expectations argue that momentum is compensation for some underlying risks not accounted for by traditional asset pricing models. Berk, Green and Naik (1999) construct a rational model based on economic risk factors that facilitates both momentum and short-term reversals. Their model generates momentum through persistence in conditional expected returns, which results from the fact that the firm's exposure to economic risk is persistent over the intermediate horizon. They suggest that the returns over the past 12 months is a good proxy for conditional expected returns because, over longer horizons, a firm's ongoing project and thus economic risk exposures change.

Johnson (2002) provides one rational interpretation of momentum based on the notion that stock prices are highly and non-linearly sensitive to firm growth rates, such that growth rate risk rises as growth rates become higher. All else equal, firms with high realized recent returns are more likely to have experienced increases in growth rates and thus more risk related to growth rates. If one assumes that investors need to be compensated for taking on this additional risk, future expected returns need to increase respectively, generating a momentum effect. Sagi and Seasholes (2007) provide a similar risk-based model. In their model, a firm can have growth options and limited liability options. Growth options are assets that increase in value, when the firm has performed well, while limited liability options reduce risk and expected returns, when the firm performs poorly. These options contribute positively to return autocovariance and the positive return autocovariance in turn gives rise to momentum.

Holden and Subrahmanyam (2002) propose a theory of rational investors with asymmetric information. In their model news arrive discretely, agents speculate on upcoming news and some agents become informed of the content of the news later than others. Uncertainty about these news gives rise to risk premia, which decrease gradually as more and more agents become informed about upcoming news. Vayanos and Woolley (2013) suggest a rational model of momentum and long-term reversals based on fund flows. They propose, that investors update their knowledge of a fund manager's efficiency through realized returns or directly. Momentum arises when outflows by rational investors cause the fund to have to liquidate part of their assets, causing a further negative shock to prices of the assets that experienced recent negative shocks. Luo (2012) posts empirical evidence consistent with this model.

2.4.3 Factor explanations

A related body of empirical work links momentum returns to different types of cross-sectional and macroeconomic factors. It is worth noting here, that there is nothing special about a factor explanation that would preclude a behavioral explanation (see e.g. Kozak, Nagel & Santosh (2018) for a recent discussion). However, most of these factors do build on the premise of rational investors and risk.

Liu and Zhang (2008) show that a little more than half of momentum returns can be explained by a priced macroeconomic risk factor constructed based on the growth rate of industrial production. Their work is motivated by the theoretical work of Johnson (2002) and Sagi and Seasholes (2007). They find that past winners temporarily experience temporary higher exposures on the growth rate of industrial production than losers. These loadings then tend to reverse to a degree after a few months.

Chordia and Shivakumar (2002) show that most of momentum returns can be explained by systematic exposures to a set of macroeconomic factors that are known to predict future market returns. While they do not explicitly provide a rational explanation, they conclude that future behavioral theories should seek to reconcile these empirically shown links to the macroeconomy. Griffin and Martin (2003) look for this relationship between momentum and macroeconomic variables around international markets and find that the conditional model of Chordia and Shivakumar (2002) does not explain momentum returns outside the US.

Pastor and Stambaugh (2003) find that about half of the returns to a momentum strategy can be explained by exposure to liquidity risk. Sadka (2006) decomposes liquidity risk into variable and fixed components and shows that the unexpected systematic variations in liquidity risk explain 40% to 80% of the returns to a momentum portfolio. Recently, Ruenzi and Weigert (2018) show that momentum has a significant positive loading on a crash sensitivity factor introduced by Chabi-Yo, Ruenzi and Weigert (2018). The factor captures investors' aversion to stocks that have large conditional betas during market crashes. They find that momentum does not generate significant positive returns after controlling for crash sensitivity risk.

Ahn, Conrad and Dittmar (2003) explore a stochastic discount factor approach that is not conditional on a particular asset pricing model being true and allows for time-varying risk premiums. They find that around half of the returns to momentum are explained by their model, concluding that momentum does not seem to be entirely consistent with rational pricing. Fong, Wong and Lean (2005) also conduct tests that are independent of any particular benchmark model. They employ a stochastic dominance approach to test whether a general asset pricing model can explain

momentum when investors are risk-averse. Like Ahn et. al (2003), they find evidence that is difficult to reconcile with rational models.

Hou, Xue & Zhang (2014) show, that their four-factor model with market, size, investment and profitability factors explain most of momentum returns. They emphasize that their model does not take a stance between risk-based and behavioral interpretations. Kelly, Moskowitz and Pruitt (2018) find that the predictive power of momentum is subsumed by a conditional factor model that aggregates return predictability from a broad range of firm characteristics known to predict returns. Kelly, Moskowitz and Pruitt (2018) posit that their factors are to be interpreted as compensation for risk. Stambaugh and Yu (2017) build on a similar premise of combining multiple characteristics into a factor model, but with a behavioral interpretation. They find that their model similarly subsumes momentum. It is worth noting that both Kelly, Moskowitz & Pruitt (2018) and Stambaugh and Yu (2017) include momentum itself in their set of characteristics.

2.5 Determinants of momentum returns

An extensive literature focuses on identifying predictable determinants of momentum returns. An extensive portion of this literature is interested in different characteristics of stocks and returns that might result in differing degrees of momentum profits. This literature is of interest in the discourse over different theories of momentum, as connections between momentum and firm-specific characteristics can shed light on how momentum arises. In addition, identifying both time-series and cross-sectional determinants of momentum returns is interesting from an investment perspective. If an investor can observe firm-specific or market-wide variables, that command higher momentum returns ex-ante, they can potentially utilize this information to improve the profitability of a momentum strategy. In this section, I cover the literature on cross-sectional and time-series determinants of momentum and limits to arbitraging momentum. Although I briefly cover some investment perspectives, my main focus is on the implications of results in terms of theories on momentum.

2.5.1 Cross-sectional determinants of momentum

First, I cover literature considering different firm and return characteristics that command higher or lower momentum profits. The key idea here is that different theories of momentum imply, directly or indirectly, that momentum should be stronger among certain types of stocks.

Da, Gurun & Warachka (2014) test an underreaction theory of momentum. Using daily returns, they show that momentum is much stronger among stocks where formation period returns accumulate more gradually. Grinblatt and Moskowitz (2004) find similar evidence using monthly returns to measure the consistency of returns instead. There are a few different interpretations for these findings. Da et al. (2014) hypothesize that investors are inattentive to news that arrive gradually, while Grinblatt and Han (2005) argue that these results are consistent with their model based on the disposition effect, because consistent winners are likely to have larger unrealized capital gains. Another potential explanation is the gradual diffusion of private information among investors, consistent with both a behavioral explanation (Hong and Stein, 1999) and a rational interpretation based on information asymmetry (Holden and Subrahmanyam, 2002).

Hong, Lim and Stein (2000) show that returns to momentum profits are markedly stronger in low-market capitalization stocks and, after controlling for market capitalization, among stocks with low analyst coverage. They interpret these findings as evidence supporting an underreaction theory of momentum. However, they could just as well be indicative of the information asymmetry and rational investors, as suggested by Holden Subrahmanyam (2002). Grinblatt and Moskowitz (2004) similarly find stronger momentum among low market capitalization stocks. However, Israel and Moskowitz (2013) show that the size effect found by Hong et al. (2000) and Grinblatt and Moskowitz (2004) is unique to the period between 1980 and 1996.

Hvidkjaer (2006) examines investor behavior by comparing trading volume initiated by buyers and sellers. He finds that small traders exhibit a strong underreaction effect followed by overreaction. Furthermore, he finds that the actions of small traders during the portfolio formation period significantly affect the profitability of momentum. He

concludes that momentum could be driven by the behavior of these small traders, consistent with a behavioral explanation. Chui, Titman and Wei (2010) hypothesize that investors in countries with high degree of individualism are more likely to be subject to the biases that behavioral theories attribute momentum to. Exploring the link between the degree of individualism and momentum returns across countries, they find that momentum is stronger in countries with high degree of individualism. They conclude that their evidence favors a behavioral interpretation of momentum.

Daniel and Titman (1999) hypothesize that mispricing caused by investor overconfidence should be stronger in stocks, whose valuation relies on more subjective information, such as growth stocks. They find that momentum returns are significantly stronger among stocks with low book-to-market ratios than stocks with high book-to-market ratios. They argue that this finding is in line with momentum caused by investor overconfidence. Sagi and Seasholes (2007) show that momentum is considerably higher among firms with high volatility in revenue growth and low cost of goods sold. They also find evidence of the book-to-market effect, initially documented by Daniel and Titman (1999). However, Sagi and Seasholes (2007) interpret their findings based on a model with rational investors instead.

Lee and Swaminathan (2000) find, that momentum is stronger among stocks with high trading volume. Furthermore, they find that strongest momentum returns are generated by strategies that go long on low-volume winners and short on high-volume losers. Hong and Stein (2007) note that the Hong and Stein (1999) model incorporating both under- and overreaction is consistent with higher momentum returns among stocks with higher trading volume.

A related finding is documented by Arena, Haggard & Yan (2008), who find that momentum is significantly stronger among stocks with higher idiosyncratic volatility. They deem their findings supportive of a behavioral explanation. They base this interpretation on the intuition that returns to stocks with higher idiosyncratic volatility have more firm-specific information and thus exhibit stronger under- or overreaction. However, Vayanos and Woolley (2013) in fact provide a rational model that predicts this relationship between momentum and idiosyncratic risk. Avramov, Chordia, Jostova and Philipov (2013) document a related finding that credit risk plays a key role

in explaining momentum. They find that, when excluding the companies with the worst credit ratings, momentum returns become statistically insignificant and economically small.

Jegadeesh and Titman (2001) find that momentum strategies earn positive returns over the first 12 months following portfolio formation, but that these returns revert to a degree over the 13 to 60 months following portfolio formation. Lee and Swaminathan (2000) show that these reversals occur faster for high volume winners and low volume losers. Both Jegadeesh and Titman (2001) and Lee and Swaminathan (2000) argue that their findings support an overreaction story. However, Jegadeesh and Titman (2001) advise caution in interpreting their results, because they are sensitive to methodological choices. Looking at both momentum and earnings momentum, Chan, Jegadeesh & Lakonishok (1996) find only very weak evidence of reversals in momentum returns and no evidence of reversals in earnings momentum. Conrad and Yavuz (2017) provide evidence that contest earlier findings of reversals. They show that the stocks, which contribute positively to momentum are in fact less likely to exhibit reversals. They find that these return reversals are driven by stocks that do not contribute to momentum returns. Furthermore, Grinblatt and Moskowitz (2004) argue that momentum and long-term reversals are at least partly unrelated, because long-term reversals and momentum exhibit quite different seasonal patterns. These findings seem to indicate that the ability to explain long-term reversals is not a particularly important prerequisite for a model of momentum.

Two recent papers study the term structure of momentum returns. Novy-Marx (2012) finds that momentum in U.S. stock returns is driven by returns over 12 to 7 months prior to portfolio formation. Momentum portfolios in the U.S. stock market, formed by using returns over the prior 6 to 2 months, significantly underperform portfolios formed on returns over the past 12 to 7 months. He argues that this echo-like term structure of momentum presents a challenge for theories, as most suggested models of momentum, both behavioral and rational, fail to deliver this type of a term structure. However, Goyal and Wahal (2015) argue that this underperformance is largely driven by a spillover of short-term reversals to month 2 prior to portfolio formation. Looking at international evidence, they find that this discrepancy between momentum

portfolios formed on months -2 to -6 and month -7 to -12, is very small or non-existent, or even reversed in major stock markets outside the U.S.

2.5.2 Momentum over time and the market cycle

Another branch of the literature documents predictability in momentum returns over time. Jegadeesh and Titman (1993) find that momentum returns exhibit a very specific and strong seasonal pattern. They find that momentum has economically large and statistically significant negative returns in Januaries. Jegadeesh and Titman (2001) confirm that this phenomenon persist also in the nine years following the Jegadeesh and Titman (1993) sample period. Grinblatt and Moskowitz (2004) hypothesize that tax-loss selling causes these seasonality effects. Consistent with this prediction, they show that these seasonal patterns are stronger among stocks that are most likely held by taxable investors and in high tax years. Grinblatt and Han (2005) show that their behavioral model of momentum can explain the December and January effects, when tax-loss selling causes negative drift in investors' demand. Sias (2007) finds that momentum returns are stronger in quarter-ending months and particularly in Decembers. Sias (2007) links these results empirically to both window dressing by institutional investors and tax-loss selling by investors.

Many papers document the pro-cyclical nature of momentum. Cooper, Gutierrez & Hameed (2004) find that momentum returns are positive only in up markets and negative in down markets. They attribute their finding to increased overreaction, stemming from an increase in investor overconfidence after positive market outcomes. Chordia and Shivakumar (2002) document the same phenomenon but propose a risk-based interpretation. Sagi and Seasholes (2007) show that their rational model generates this pro-cyclicality as well. Antoniou, Doukas and Subrahmanyam (2013) argue for a behavioral interpretation and show that investor sentiment has incremental value in explaining momentum returns on top of market returns. Stivers and Sun (2010) explore cross-sectional return dispersion as a state variable and find that momentum returns are lower following high cross-sectional return dispersion and vice versa. They further show that the explanatory value of past market returns is subsumed by the explanatory value of return dispersion.

Daniel and Moskowitz (2016) examine crashes in momentum returns. They find that momentum strategies experience occasional large negative returns commonly occurring after the stock market recovers from downturns, because momentum tends to have a large negative conditional market beta during these times. Earlier work by Grundy and Martin (2001) documents this market state dependent beta of momentum as well. The cyclical nature in momentum returns found in earlier studies is potentially explained by these occasional crashes, although Daniel and Moskowitz (2016) do not directly investigate this. Daniel and Moskowitz (2016) explore potential explanations for their findings based on risk but find no explanation that can fully account for their results. They conclude that their findings could be consistent with either a behavioral or a rational explanation. The findings of Daniel and Moskowitz (2016) are in conflict with the findings of Ruenzi and Weigert (2018), who show that exposure to a risk factor based on investors' aversion to tail risk explains most of momentum profits. Because momentum in fact takes on short positions in stocks that are sensitive to negative tail-events, it seems counterintuitive that momentum would be compensation for holding tail-risk sensitive assets.

Barroso and Santa-Clara (2015) examine the tail-risk of momentum strategies and suggest an improved momentum strategy that scales its long-short position based on the realized volatility of momentum returns. They find that this adjustment almost doubles the Sharpe ratio of momentum. Daniel and Moskowitz (2016) show that taking market state into account further improves upon the volatility-managed momentum strategies suggested by Barroso and Santa-Clara (2015).

2.5.3 Momentum and limits to arbitrage

Because momentum is a widely documented phenomenon, there needs to exist some reason why momentum returns continue to persist. If momentum returns are reward for some form of risk taken on by momentum investors, then this reconciliation is straightforward. However, if momentum is caused by behavioral biases, then rational investors can take advantage of these biases by employing momentum strategies. As more rational investors engage in harvesting momentum returns, momentum should attenuate. If momentum is to persist, investors need to remain irrational despite being aware of momentum existing. This bias needs to be either widespread or large in

magnitude in order to exceed the extent, to which rational arbitrageurs are able to take advantage of momentum returns. The ability of rational arbitrageurs to take advantage of momentum is limited by trading costs and other limits to arbitrage.

Lesmond, Schill and Zhou (2004) estimate the trading costs related to momentum, using various methods suggested in the literature. They find that momentum strategies have high turnover, and the stocks driving momentum tend to have high transaction costs. Korajczyk and Sadka (2004) document similar results, estimated with alternative measures. Frazzini, Israel and Moskowitz (2012), however, argue that preceding literature exaggerates true trading costs that a large institutional investor might face. Using data of realized transaction costs including price impact, incurred by a large asset management firm, they show that true realized transaction costs are around a tenth of the costs suggested by previous literature. They find that momentum is robust to these trading costs and can be implemented on a very large scale.

Shleifer and Vishny (1997) and Pontiff (2006) identify idiosyncratic volatility as an important factor limiting arbitrage. McLean (2010) focuses on the arbitrage limiting aspect of idiosyncratic volatility and finds no significant relationship between idiosyncratic volatility and momentum. He suggests that arbitraging momentum is limited by high transaction costs instead. The findings of McLean (2010) contradict those of Arena, Haggard & Yan (2008). The key difference between the two papers is in their methodological approach. McLean (2010) emphasizes the limits to arbitrage – aspect of idiosyncratic volatility and examines portfolio selection by a risk-averse investor, while Arena et al. (2008) study the magnitude of momentum returns across terciles of idiosyncratic volatility.

2.6 Momentum and return autocovariance

The set of explanations offered for the momentum premium is large. Subrahmanyam (2018) calls for more research seeking to separate between the theories. However, the literature discussed in subsections 2.3-2.4 demonstrates the difficulty of this task. Models of momentum generate a variety of testable predictions and empirical evidence exists in support of many different interpretations. One important question to assess when evaluating different theories is what return mechanisms drive momentum. Lo

and MacKinlay (1990) show that returns to strategies based on past returns can be decomposed three different sources of returns: return autocovariances, return cross-covariances and cross-sectional variation in unconditional expected returns. Lewellen (2002) provides a thorough description of how these different return drivers are related to theories on momentum. He also submits empirical evidence that is difficult for most models in the literature to reconcile.

Throughout the remainder of the thesis, I refer to return autocovariances and cross-covariances. Many papers on decomposing momentum elect to use return autocorrelation and cross-correlation instead. In terms of theories on momentum, we are most interested in the sign and significance of the time-series dependence, in which case correlation and covariance can be used interchangeably. My choice of using autocovariances and cross-covariances is motivated by the fact that computing the decomposition in terms of return autocovariances and cross-covariances allows us to directly interpret decomposition estimates as the investment return generated by each return mechanism.

2.6.1 The decomposition and theories on momentum

Lo and MacKinlay (1990) examine strategies that weight assets relative to their past returns compared to an equally-weighted index of all assets. They analytically solve for the expected returns to such a strategy and uncover three distinct components.⁵ In the context of momentum these different components can be interpreted as follows. First, stock returns might exhibit positive autocovariances, such that a stock's past return positively predicts returns in following periods. Second, stock returns might exhibit negative cross-covariances. This means that a high return on one stock might negatively predict subsequent returns on some other stocks. Third, some stocks might simply have higher unconditional expected returns than others. A momentum strategy on average invests long into assets with higher expected returns and short into assets

⁵ The decomposition is discussed in more detail in Section 4, when I present my data and methodology.

with low expected returns. If the dispersion in unconditional expected returns is high enough, this gives rise to momentum returns.⁶

Lewellen (2002) shows how these three return drivers are linked to models of returns under which momentum can arise. The first model considered by Lewellen (2002) assumes that prices follow a random walk, with auto- and cross-covariances being zero. In this case momentum returns are solely based on the cross-sectional variance in expected returns, and the realized past return i.e. momentum measure only provides a very noisy proxy of the unconditional expected return. In terms of theories, momentum arising from differences in unconditional expected returns would be the easiest to reconcile based on rational pricing. The intuition would be that some firms have higher unconditional returns on average and consistently end up as winners, because these firms are perceived as riskier for some reason or another. This is of course not to say that cross-sectional variation in expected returns driving momentum would completely exclude behavioral theories either. Investors might irrationally prefer holding some types of stocks over others, causing consistent differences in returns.

A common interpretation of momentum is that it is a result of investors underreacting to new information. Lewellen (2002) formally shows that when prices do not completely accommodate new information, return volatility decreases and returns become positively autocorrelated. Linking this result back to the Lo and MacKinlay (1990) decomposition, he shows that underreaction leads to momentum. This mechanism illustrated by Lewellen (2002) is in alignment with most of the behavioral underreaction theories suggested in the literature.

Lewellen (2002) also illustrates a less common interpretation that momentum arises from negative cross-covariances among stocks. He presents two models under which cross-covariances might explain momentum. The first one is a behavioral model

⁶ Lo and MacKinlay (1990) are interested in short-term contrarian strategies, but their decomposition applies for all past return strategies, including momentum. More precisely, contrarian strategies assign negative (positive) weights to assets with positive past returns. The interpretation changes in that for contrarian strategies *negative* autocovariances contribute positively, *positive* cross-covariances contribute positively and unconditional expected returns always contribute negatively.

related to overreaction where investors receive news about one firm and overestimate the effect of these news on other firms. Lewellen (2002) shows that such a model induces positive covariances in contemporaneous returns, but negative auto- and cross-covariances. Linking this back to the momentum decomposition, momentum returns are positive, if the negative cross-covariances dominate the negative autocovariances, which work against momentum. Lewellen (2002) shows that a simple restriction, that positive news on one firm have a positive but smaller effect on other stocks, results in positive momentum returns.

Lewellen (2002) shows that momentum from negative cross-covariances can alternatively be a result of time-varying risk premiums, when price changes are driven by both changes in risk premiums and cash flows. He shows that positive momentum profits arise when stocks, whose prices are sensitive to changes in risk premiums, also have cash flows with low sensitivity to changes in risk premiums and vice versa. This setting can arise, for example, because of shorter (longer) durations in the cash flows of stocks that are sensitive (insensitive) to business conditions. Lewellen (2002) notes that, since both overreaction and time-varying risk premiums generate similar patterns in the Lo and MacKinlay (1990) decomposition, distinction between these two models, where momentum is generated by negative cross-covariances, is difficult.

Lewellen (2002) illustrates these models as only some examples of the potential mechanisms driving momentum. The intuition in Lewellen (2002) seems to be, however, that positive autocovariances would be most consistent with an underreaction theory of momentum. In contrast to this, many of the rational models suggested in the literature, in fact, imply positive return autocovariances. For example, Johnson (2002), Holden and Subrahmanyam (2002) and Sagi and Seasholes (2007) all present a risk-based theory for momentum that is entirely consistent with positive return autocovariances. Similarly, most overreaction theories presented in the literature imply positive autocovariances that later turn into negative ones, instead of the kind of overreaction suggested by Lewellen (2002) above. Exploring the literature, I find no well-known theory of momentum that suggests negative cross-covariances, rational or behavioral. Furthermore, many risk-based explanations for momentum explore time-varying, *conditional* expected returns instead of *unconditional* expected returns (Berk, Green & Naik, 1999, Chordia & Shivakumar, 2002 and Kelly,

Moskowitz & Pruitt 2018). It is difficult to infer how momentum arising in a conditional setting would manifest itself in terms of the unconditional decomposition.

This is not to say that the Lo and MacKinlay (1990) decomposition is not useful in distinguishing between theories. It still provides a solid framework for disentangling the drivers of momentum returns. The most important insight of the decomposition is that momentum is not necessarily driven by positive autocovariances. Return autocovariance is a time-series attribute of an individual stock, while cross-sectional momentum is about relative returns between stocks. However, most theories seem to view evidence of cross-sectional momentum as evidence of positive autocovariance in returns and attempt to explain it as such. Now, if positive autocovariance is a prerequisite for a model, we intuitively should reject the model if we find that momentum is not driven by positive autocovariances. Instead, we should then pursue to establish models where momentum arises from one or both of the remaining drivers of momentum.

2.6.2 Empirical evidence on the decomposition of momentum

Having introduced the theoretical framework, I now summarize the empirical findings in the literature that builds upon the decomposition by Lo and MacKinlay (1990). At this point, it is worth noting that a slightly problematic facet of this branch of research is the inconsistency of the strategies employed. Some papers study momentum in individual stocks, others interpret portfolio-level momentum to be representative of both effects, based on Grinblatt and Moskowitz (1999) and Lewellen (2002). Furthermore, the choice of formation and holding period lengths varies a lot between the papers.

Conrad and Kaul (1998) decompose the returns to momentum strategies on individual stocks. They construct similar relative weighting strategies and utilize the Lo and MacKinlay (1990) framework. However, their decomposition diverges slightly from the one presented above. Conrad and Kaul (1998) only consider autocovariances and cross-covariances through their joint contribution to the returns of a momentum strategy and do not disentangle the effects of firm-specific return autocovariance and cross-covariance. They find that, for momentum strategies formed based on past 12

month returns (not skipping the latest month), the time-series predictability component contributes negatively with high statistical significance, while cross-sectional variation in unconditional expected returns is very large and positive, resulting in positive and statistically significant momentum returns. Their findings are similar for momentum strategies formed on 3 to 9 months of past returns and consistent across time. Conrad and Kaul (1998) further validate their results based on bootstrap simulations.

Jegadeesh and Titman (2001) discuss Conrad and Kaul's (1998) findings and show that momentum returns tend to revert over time, which is inconsistent with momentum arising from variation in unconditional expected returns. This finding of reversals in momentum returns, however, has been later contested by Conrad and Yavuz (2017). A more important shortcoming of Conrad and Kaul (1998) is pointed out by Jegadeesh and Titman (2002). They show that estimates of cross-sectional variance in unconditional expected returns pick up not only the variance in true unconditional expected returns but also the variance generated by estimation error. This leads to exaggerated estimates of cross-sectional variances in returns, which is further exacerbated by the fact that Conrad and Kaul (1998) include stocks with short return histories in their sample. They also show that Conrad and Kaul's (1998) bootstrap results suffer from the same small-sample bias than the estimates. This is because their method of sampling with replacement causes extreme returns to contribute heavily to momentum returns.

Park and Kim (2014) replicate the decomposition used by Conrad and Kaul (1998) in 14 different markets. They find, that momentum returns are overwhelmingly driven by variation in unconditional expected returns, while the contribution of the time-series components is in general negative. Park and Kim (2014), however, do not describe their estimations of the components in detail and do not comment on the critique presented in response to Conrad and Kaul (1998). The problems highlighted by Jegadeesh and Titman (2002) can only be exacerbated by the fact that Park and Kim (2014) study markets with shorter return histories and fewer stocks.

Bulkley and Nawosah (2009) suggest another test of whether unconditional expected returns can explain momentum. They argue that if momentum arises from variation in unconditional expected returns, then momentum should not be present when firm-

specific unconditional expected return is deducted from returns. Bulkley and Nawosah (2009) use lifetime mean return of a stock as a proxy for the unconditional expected return and find no momentum in demeaned returns. Bhootra (2011) shows that the results in Bulkley and Nawosah (2009) can, to an extent, be attributed to their inclusion of penny stocks and the fact that they elect to not skip the latest month in portfolio formation. According to Bhootra (2011), controlling for these methodological oddities makes momentum robust in demeaned returns as well. Still, returns to momentum based on demeaned returns are markedly lower than for regular momentum. The more significant shortcoming of Bulkley and Nawosah (2009) is the fact that they use all of the return observations for each asset to form their estimates of unconditional expected returns, resulting in a similar bias as in Conrad and Kaul (1998). Jegadeesh and Titman (2002) show that momentum strategies are equally, if not more, profitable when implemented on abnormal returns. Abnormal returns here refer to returns calculated as the returns over the formation period adjusted for an *out-of-sample* estimate of the unconditional expected return.

Lewellen (2002) studies the full decomposition of momentum returns in industry portfolios and Fama and French size and book-to-market sorted portfolios. He considers relative weighting strategies that are formed based on 12-month returns and held for one month with zero to 17 months skipped between the formation period and the holding period. The results Lewellen (2002) documents are striking. He finds that, across the 18 lags, both auto- and cross-covariances in returns are mostly negative. These results are consistent across size-sorted portfolios, book-to-market-sorted portfolios and industry portfolios with around half of the negative estimates being statistically significant. Lewellen (2002) posits based on his results, that momentum must be driven by return cross-covariances.

Pan, Liano and Huang (2004) focus on similar strategies but using weekly returns. They argue that the resulting larger sample size allows for more powerful tests. Pan et al. (2004) look at return autocovariances of 20 industry portfolios using weekly returns at lags of 1 to 26 weeks. They find positive autocovariances on short horizons, which gradually decrease and turn into negative ones at horizons longer than 12 weeks. The autocovariances at longer lags are much smaller in magnitude. Turning to the momentum strategies, Pan et al. (2004) form relative weighting strategies based on

one-week returns at varying lags and held for one week. They find that momentum is very strong and significant on short horizons, but strategies formed on returns in the past 12-26 weeks generate no returns. Decomposing the returns, they find that positive autocovariances drive this momentum effect, cross-covariances contribute significantly negatively and the unconditional expected returns have a very small but statistically significant positive contribution. These findings are in stark contrast to Lewellen's (2002) findings on monthly returns. Differences are to be expected due to the major deviations between the methodologies in the two papers. Because weekly returns are very noisy, it is unlikely that there is enough signal relative to noise at lags longer than a few weeks to form a meaningful investment strategy. Moreover, it is well known that individual stock returns exhibit reversals at shorter horizons, so it is unlikely that these positive autocovariances at short horizons in portfolios are meaningful in terms of explaining momentum in individual stocks.

Moskowitz, Ooi, Pedersen (2012) find that positive autocovariances drive momentum in equity indices, commodities, government bonds and currencies. Furthermore, they find that time-series momentum is stronger than cross-sectional momentum in these asset classes. Moskowitz et al. (2012) argue that returns to time-series momentum are higher, because time-series momentum only picks up return autocovariances, which seem to contribute positively to momentum, and ignores cross-covariances that contribute negatively. Goyal and Jegadeesh (2018), however, note that the better performance of time-series momentum strategies is due to a time-varying net long position in the equal-weighted market index that the time-series momentum strategy takes on. Thus, they argue that returns to time-series momentum are not directly comparable to cross-sectional momentum, which is a zero net investment strategy.

Holden and Subrahmanyam (2002) study autocovariances of decile portfolios sorted by market capitalization and past return over the last 3-months, to study their information asymmetry theory for momentum. They find positive autocovariances among all past return deciles and all size deciles except for the largest decile of stocks. The positive autocovariances are statistically significant in the 5 smallest deciles of market capitalization and in 8 out of the ten past return deciles. Holden and Subrahmanyam (2002) argue that the short 3-month horizon is most relevant to their study, because their model is focused on earnings announcements and similar news,

which are released quarterly. They mention that using a return horizon of 6 and 12 months yields similar but weaker results. Exact results are, however, left unreported. Holden and Subrahmanyam (2002) do not relate their findings to other literature or directly to the decomposition of momentum returns.

Pan (2010) constructs momentum strategies on industry and characteristic-sorted portfolios, following Lewellen (2002). Using formation and holding periods of equal length and varying from 3 to 12 months, they find evidence supporting cross-covariances as the key driver of momentum, similar to Lewellen (2002). However, Pan (2010) argues that since the relevant return horizon in Lo and MacKinlay (1990) is fixed, the decomposition fails to capture positive autocovariances at shorter horizons, which is the type of autocovariances predicted by behavioral models. To analyze this intuition empirically, Pan (2010) examines monthly autocovariances across 12 lags. He finds that the average and cumulative autocovariances are positive across all industry and characteristic-sorted portfolios over a 6-month horizon and positive for most portfolios over a 12-month horizon. Based on this, Pan (2010) concludes that returns exhibit the kind of positive autocovariance that is consistent with most underreaction theories. However, Lo and MacKinlay (1990) decomposition is a mathematical identity that unambiguously defines the relevant return horizon for return autocovariance. Pan (2010) does not discuss any mechanism that would cause a discrepancy between the relevant horizon for momentum and return autocovariance.

Chen and Hong (2002) show that the Lo and MacKinlay (1990) can yield inaccurate results when returns follow a common factor structure and the factor itself exhibits autocovariance. To account for this shortcoming of the decomposition, Chen and Hong (2002) present an alternative decomposition following Jegadeesh and Titman (1995). This decomposition relies on the assumption that returns follow a common factor structure with constant exposures to contemporaneous and lagged realizations of the common factor. They study momentum in industry portfolios, using 6-month formation and holding periods, with no months skipped in between. They find that, in this alternative decomposition, positive return autocovariances show up as the dominant determinant of momentum returns, being responsible for around 80% of the momentum returns in their full sample from 1928 to 1999. The overreaction component, consisting of the contemporaneous and lagged factor exposures,

contributes positively but account for less than 20% of the momentum returns in the full sample. Finally, Chen and Hong (2002) find that the contribution of variation in unconditional expected returns is only around 5% of the total returns.

Du (2012) employs a regression-based methodology, proposed by Du and Watkins (2007), that allows for decomposing momentum returns into all three components in individual stocks. The strategies he employs are portfolio-based momentum strategies, with one-month and 6-month formation period and a 6-month holding period. His results are somewhat similar to the findings of Lewellen (2002). He finds that over both horizons considered, firm-specific autocovariances contribute negatively and statistically significantly to momentum returns. He also finds that variation in unconditional expected returns and negative cross-autocovariances contribute about equally to the 1-month strategy. For the 6-month formation period, variation in unconditional expected returns arises as the most important component.

3 RESEARCH QUESTIONS AND CONTRIBUTION TO THE LITERATURE

Lo and MacKinlay (1990) decompose returns to past return strategies and show, that returns to a momentum strategy can be caused by either positive return autocovariances, negative return cross-covariances or by cross-sectional variation in unconditional expected returns. In subsection 2.4, I cover a large body of theoretical literature suggesting varying interpretations of the momentum anomaly, including both rational and behavioral interpretations. In subsection 2.6, I discuss, how understanding the return mechanism behind momentum is important in distinguishing between these theories. All behavioral, and some rational theories on momentum imply that momentum is driven by positive autocovariances in returns. Some rational theories suggest that momentum is caused by cross-sectional variation in unconditional expected returns. Lewellen (2002) suggests both a rational and a behavioral interpretation of momentum, when momentum is driven by negative cross-covariances among stocks. However, prior empirical evidence on which of these return mechanisms are important in explaining momentum is inconsistent. Empirical evidence in support of all three return mechanisms has been posted in the literature.

To summarize, we have many competing theories on momentum but inconclusive empirical evidence on the return mechanism behind momentum. In my empirical analysis, I add to this branch literature on decomposing momentum strategies. Explicitly stated, I focus on the research question: Is momentum driven by positive return autocovariances, negative cross-covariances or cross-sectional variation in unconditional expected returns? For each of these three return mechanisms, I separately test for the hypothesis that the component positively contributes to the returns on momentum strategies.

I contribute to prior literature by examining the decomposition in both individual stocks and portfolios of stocks. Prior work on individual stocks focuses on the role of unconditional expected returns and does not separate between autocovariances and cross-covariances. On the other hand, prior research that does disentangle autocovariances from cross-covariances focuses on portfolios of stocks instead of individual stocks. Examining momentum in individual stocks and portfolios separately

is important because momentum in portfolios of stocks and momentum in individual stocks might be somewhat separate phenomena. For example, Grundy and Martin (2001), Chordia and Shivakumar (2002) and Blitz, Huij & Martens (2011) present evidence that momentum in stock portfolios and momentum in individual stocks are somewhat distinct phenomena. Furthermore, return auto- and cross-covariances can behave differently for individual stocks and portfolios of stocks. For example, Demiguel, Nogales & Uppal (2014) find that, in *daily returns*, first-order autocovariances and cross-covariances are negative for individual stocks while they are positive for portfolios of stocks. Motivated by this prior literature, I seek to answer a secondary research question: Do the three return drivers contribute differently to momentum in individual stocks and to momentum in portfolios of stocks?

There are likely two reasons why prior literature focuses on portfolios of stocks instead of individual stocks. First, estimation is challenging for individual stocks, because of unbalanced panels of return histories. The second problem with using individual stocks is that samples of individual stocks include extreme return observations and short return histories. This causes extreme return observations to disproportionately contribute to sample estimates of autocovariances, cross-covariances and unconditional expected returns, resulting in biases in the decomposition.

I suggest two complementing methodologies that allow me to fully decompose the returns for both individual stocks and at the portfolio level. The first one involves conducting the decomposition in a moving window, using pairwise return histories. This enables me to estimate cross-covariance matrices for individual stocks in the presence of unbalanced panels. The second method I suggest is a strategy-based decomposition, novel to the literature. I analytically decompose the momentum strategy into three separate strategies that take advantage of autocovariances, cross-covariances and unconditional expected returns respectively. This allows me to restate the question of which return mechanism drives momentum as: which return mechanisms can an investor take advantage of in real time? Because the strategy-based decomposition does not require estimating cross-covariance matrices, it does not restrict the universe of stocks to those with pairwise return histories.

Jegadeesh and Titman (2002) demonstrate, that estimates of the variation of unconditional expected returns are positively biased for individual stocks. Statistical literature (Chan, 1992, Chan, 1995 and Deutsch, Richards & Swain, 1990) documents the sensitivity of sample estimates of autocovariances and cross-covariances to outlier observations. My third main research question is motivated by this literature and the inconsistency of prior evidence on decomposing momentum returns. My third question, explicitly stated is: Do sample estimates of the contributions of the different return mechanisms, used in prior literature suffer from biases? To test for these biases, I replicate a bootstrap experiment, used by Jegadeesh and Titman (2002) to study individual stocks, and extend their evidence to portfolio data. I utilize the same bootstrap setting and make a novel contribution to the literature by showing that sample estimates of the contribution of return autocovariances are negatively biased and sample estimates of the contribution of cross-covariances are positively biased. These biases likely contribute to the inconsistencies in prior empirical evidence.

I tackle these biases in the decomposition in two ways. I provide an initial examination by estimating the decomposition using winsorized returns, pulling in the extreme returns. My main method of adjusting for the biases is the strategy-based decomposition. The strategy-based decomposition gives a natural interpretation to these biases as forward-looking biases. Using only information available to the investor ex-ante to form these strategies results in unbiased estimates of the decomposition. I show that both methods of controlling for these biases result in positive autocovariances appearing as the key driver of momentum returns.

I further contribute to the literature by examining, how these contributions evolve over time. Prior literature generally assumes, that the contributions of autocovariances, cross-covariances and cross-sectional variances in unconditional expected returns remain consistent throughout the sample period. Another advantage of the strategy-based decomposition is that it naturally allows for time-variance in the contributions. Lo and MacKinlay (1990) show that the accuracy of their decomposition does not rely on the assumption of stationarity and that estimates of each contribution can be interpreted as the average contributions. However, the return mechanisms behind momentum may change over time, which is quite relevant with regards to theories on

momentum. The time-evolution of these components also sheds some light on the potential profitability of momentum in the future.

My final contribution to the literature is to extend the sample period under study. Prior literature on the decomposition mostly dates back to the early 2000s. I provide new evidence on both portfolio- and stock-level momentum, including during and after the financial crisis of 2007-2008, a period during which momentum has struggled.

4 DATA & METHODS

4.1 Data

My primary data consists of monthly returns on individual U.S. stock returns from the Center for Research in Security Prices (CRSP) from January 1926 to the June of 2018. I focus on U.S. common stocks (stocks with the share code 10 or 11 in the CRSP database). In addition, monthly market capitalization data from the CRSP database is collected. The sample consists of 3,143,667 firm-month observations, after forming momentum, which requires at least 13 months of return history.

Microcap stocks represent only a very small fraction of total market capitalization and can be influential on results (Fama & French, 2008). Because of this, I further restrict the sample to exclude microcap stocks for my main results. I follow Fama and French (2008) and define microcaps as stocks below the 20th percentile of market capitalization. I also exclude stocks with missing market capitalization data. This restricts the sample into 2,514,738 firm-month observations. I reintroduce microcaps to the sample later in robustness tests. In robustness checks, I also consider megacaps, defined as the stocks above 80th percentile of market capitalization, large stocks, defined as the stocks between the 50th and 80th percentile of market capitalization, and small stocks, defined as the stocks between the 20th and 50th percentile of market capitalization, separately.

In addition to the individual stocks, I decompose momentum in value-weighted industry and characteristic-sorted portfolios. This allows for making some meaningful comparisons to prior literature on momentum decomposition that mainly focuses on portfolios. Four portfolio datasets are considered. The first two are 6 and 25 Fama and French value-weighted portfolios, double sorted by book-to-market ratio and market capitalization, henceforth referred to as Fama and French portfolios. The other two portfolio datasets are 10 and 30 value-weighted portfolios, sorted by industry, henceforth referred to as industry portfolios. The monthly return data for all portfolio datasets are collected from Kenneth French's data library (French, 2018).

Panel A of table 1 presents the summary statistics for market capitalization, returns, momentum and the length of the return series, after the formation of momentum, which takes up 13 return months. Panel C of table 1 presents the correlations between these variables for the CRSPs dataset, after the formation of momentum, with Spearman rank-order correlations in above-diagonal and Pearson product-moment correlations in below-diagonal entries. The return length column presents summary statistics for the length of stocks return histories. We see from Panel A of table 1 that, after the formation of momentum, the average return history is only 82 return history months, and the sample contains stocks with just 14 return observations.

Table 1. Summary statistics of the individual stock and portfolio data samples from 1926 to 2018.

Panel A: Summary statistics for individual stock data				
	MV	Ret	Mom	Return length
Min	0.459	-0.981	-0.998	14
Mean	231.198	0.011	0.136	82
Median	114.697	0.003	0.071	138.5
Max	999.998	10.344	49.980	1102
SD	258.845	0.143	0.571	154.2
Panel B: Return statistics for the portfolio data				
	FF6	FF25	Ind10	Ind30
Min	-0.351	-0.495	-0.348	-0.509
Mean	0.011	0.012	0.010	0.010
Median	0.013	0.013	0.011	0.011
Max	0.836	1.475	0.798	1.254
SD	0.068	0.077	0.060	0.070
Panel C: Correlations of the key variables in the CRSP dataset				
	MV	Ret	Mom	Return length
MV		0.022	0.18	0.223
Ret	-0.008		0.05	0.045
Mom	0.109	0.014		0.109
Return length	0.205	0.014	0.027	

Panel A contains statistics for the four key variables in the CRSP sample of individual stocks. MV = market capitalization in 1 000 000s, Ret = monthly return, Mom = the cumulative 11-month past return one month prior to the return month, Return length = length of the return series after the formation of the momentum variable.

Min = sample minimum, Mean = sample average, Median = sample median, Max = sample maximum, SD = standard deviation.

Panel B contain summary statistics for the returns of the four portfolio datasets. FF6 = Six market capitalization and book-to-market sorted portfolios, FF25 = 25 market capitalization and book-to-market sorted portfolios, Ind10 = 10 industry portfolios, Ind30 = 30 industry portfolios.

In panel C, below-diagonal entries represent Pearson product-moment correlations and above-diagonal entries represent Spearman rank-order correlations.

The correlations in Panel C of table 1 reveal that return length has a positive correlation with market capitalization, returns and momentum. The high positive correlation of between market capitalization and return length suggests, that the short return history stocks are mostly smaller stocks. The positive correlations between return length and returns and momentum, suggest that these short return history stocks also earn lower returns on average.

4.2 Relative-weight momentum and the Lo and MacKinlay (1990) decomposition

Prior work on decomposing momentum varies in their use of formation and holding periods. I focus is on the most common measure used in the literature, the cumulative returns of an asset compounded over months $t - 12$ through month $t - 2$. One month is skipped between the portfolio holding period and the formation period to avoid the reversals that returns tend to exhibit at shorter horizons (Jegadeesh, 1990, Lehmann, 1990). For momentum in portfolios of stocks, the month before portfolio holding period has a significant positive effect on performance instead (Grundy & Martin, 2001). In robustness checks, I consider strategies that do not skip the latest month for all datasets, as well as alternative formation and holding periods in subsection 5.6.

The strategies I consider are *relative strength strategies* that hold assets in proportion to their return at time $t - 2$ relative to the return of an equal-weighted index. The weights $w_{i,t}$, for asset i in month t is given by

$$w_{i,t} = \frac{1}{N} (r_{i,t-2}^{11} - \bar{r}_{t-2}^{11}), \quad (1)$$

where $r_{i,t-2}^{11}$ is the realized cumulative 11-month return of an asset at month $t-2$, i.e.

$$r_{i,t-2}^{11} = \prod_{t-13:t-2} (1 + r_{it}) - 1. \quad (2)$$

The term \bar{r}_{t-2}^{11} in equation (1) is the equally weighted mean of the 11-month returns on all assets at $t-2$. The weights add up to zero, meaning that the strategy is market-neutral long/short strategy, and the weights are scaled each period such, that each month, the strategy invests 1\$ long and 1\$ short. The return on the portfolio in period t is then

$$\pi_t = \frac{1}{N} \sum_{i=1}^N r_{i,t} (r_{i,t-2}^{11} - \bar{r}_{t-2}^{11}). \quad (3)$$

Lo and MacKinlay(1990) show that rearranging equation (3), and taking expectations on both sides yields:⁷

$$\begin{aligned} E[\pi_t] &= \frac{1}{N} E \left[\sum_i r_{i,t-2}^{11} r_{i,t} \right] - \frac{1}{N} E \left[\bar{r}_{t-2}^{11} \sum_i r_{i,t} \right] \\ &= \frac{1}{N} \sum_i (\rho_i + \mu_i^2) - (\rho_m + \mu_m^2) \end{aligned} \quad (4)$$

Where ρ_i and ρ_m are the respective autocovariances (between 1-month and 11-month returns) of asset i and the equally-weighted market index, and μ_i and μ_m are their unconditional expected returns. Equation (4) shows, that the returns to momentum depend on the assets' autocovariances relative to the autocovariance of the equally

⁷ The full derivation of the decomposition is shown in Appendix 1.1

weighted market index. Rearranging this equation, we can write the expected return in matrix notation as

$$E[\pi_t] = \frac{1}{N} \text{tr}(\Omega) - \frac{1}{N^2} \mathbf{1}' \Omega \mathbf{1} + \sigma_{\mu\gamma}. \quad (5)$$

In equation (5), $\mathbf{1}$ denotes an $(N \times 1)$ vector of ones, Ω is the autocovariance matrix $\Omega \equiv E[(r_{t-2}^{11} - \gamma)(r_t - \mu)']$ and the notation $\text{tr}(\cdot)$ denotes the trace of a matrix. In equation (5), the first term is the cross-sectional average of the autocovariances of the individual securities, the second term is the autocovariance of the equal-weighted market index, and the third term $\sigma_{\mu\gamma}$ is the cross-sectional covariance between the unconditional 11-month expected returns, $\gamma = E[r_{i,t}^{11}]$, and the unconditional 1-month expected returns $E[r_{i,t}]$. The term $\sigma_{\mu\gamma}$ represents the contribution of variation in unconditional expected returns. Equation (5) can be further rearranged into:

$$E[\pi_t] = \frac{N-1}{N^2} \text{tr}(\Omega) - \frac{1}{N^2} [\mathbf{1}' \Omega \mathbf{1} - \text{tr}(\Omega)] + \sigma_{\mu\gamma}. \quad (6)$$

Now, if we denote $C \equiv -\frac{1}{N^2} [\mathbf{1}' \Omega \mathbf{1} - \text{tr}(\Omega)]$ and $O \equiv \left(\frac{N-1}{N^2}\right) \text{tr}(\Omega)$, equation (6) can be rewritten as:

$$E[\pi_t] = O + C + \sigma_{\mu\gamma}. \quad (7)$$

In equation (7), C represents the contribution of cross-covariances to the returns of the momentum strategy, O represents the contribution of stock-specific autocovariances and $\sigma_{\mu\gamma}$ represents the contribution of cross-sectional variation in unconditional expected returns. The term O depends only on the diagonal elements of the

autocovariance matrix, i.e., the own-autocovariances in returns of each asset and the term C only depends on the off-diagonal elements of the autocovariance matrix, i.e. assets' cross-covariances. The negative sign on C and positive sign on O show that positive cross-covariances contribute negatively to momentum, while positive autocovariances contribute positively, and vice versa.

4.3 Decomposition estimation

I begin my empirical analysis with an examination of the standard decomposition used in prior literature. I estimate momentum returns and the decomposition for the stock- and portfolio-level datasets. I use sample auto- and cross-covariance estimates to form the cross-covariance matrix Ω . This is straightforward for portfolio-level data, but difficult for individual stocks, because we have a large number of assets, which all have varying lengths of return histories, including some quite short time-series of return observation. I tackle this problem by estimating the decomposition with a rolling-window methodology as follows. For every 12-month period in the sample, I find the stocks that have full return histories over the past 24 months. This is done to ensure that our set of stocks has at least 24 months of overlapping return observations. I then estimate the momentum decomposition on these selected stocks over the entire sample period. I estimate the autocovariances using the full time-series' of returns over the entire sample period, and cross-covariances using full overlapping pairwise return histories for each pair of sampled stocks. Moving forward 12 months in time, and repeating the above procedure allows us to form an annual time-series of component estimates.

The requirement for 24 months is a tradeoff between estimation error and data availability. For individual stocks, we need a period of where all sampled assets have non-missing returns. Requiring 24 months excludes stocks with shorter return histories than 36 months, because we require 12 months initially to form momentum. A longer estimation period eliminates stocks with short return histories from the sample, introducing a survivorship bias. Shorter horizon introduces more noise relative to signal in the estimation of the contributions. Furthermore, short return histories cause positive bias in estimates of the cross-sectional variation in unconditional expected returns, because unconditional expected returns are estimated with error, and the

estimate for the variation in unconditional expected returns picks up this error in addition to the true variation.

While this method results in an annual time-series of the contributions, the time-evolution in the time-series of estimated components only result from evolution in the universe of available stocks, because auto-covariances, cross-covariances and unconditional expected returns are estimated using full return histories. To examine time-variation in the contributions, I utilize a strategy-based decomposition instead. The strategy-based decomposition is discussed in detail in subsection 4.5.

It is important to note that scaling the momentum strategy weights creates a distinction between the momentum strategy and the components in the decomposition. The investment weights in the momentum strategy are scaled such that they 1\$ long and 1\$ short each period. This means that the weights are multiplied by a scaling factor, $s_t = 2 * \sum_{i=1}^N \frac{1}{|r_{i,t-2}^{11} - \bar{r}_{t-2}^{11}|}$, which is a multiple of the inverse of the cross-sectional mean absolute deviation of the returns over the momentum period. This introduces a conditioning on cross-sectional return dispersion to momentum. Now, if momentum returns vary conditional on the dispersion in returns over the momentum, the unconditional decomposition is incompatible with the actual momentum returns. Prior literature (see e.g. Lewellen, 2002 and DeMiguel, Nogales & Uppal, 2014) chooses to ignore this inconsistency and treats s , as a constant, using the expected value of s , $E[s] = \frac{1}{T} \sum_{t=1}^T s_t$, to scale the contributions to the right magnitude. I follow this convention in estimating the standard decomposition.

Statistical significance for the estimated contributions is difficult to assess analytically. For the portfolio data, I follow Lewellen (2002), and measure statistical significance of the components using bootstrapped critical values. I first demean the return series and then sample individual returns for each asset from the demeaned return series with replacement. I then compute the decomposition for a momentum strategy using these artificial time-series of returns. Repeating this procedure 500 times results in a bootstrap distribution of component estimates under the null hypothesis of no time-series predictability and no cross-sectional variation in unconditional expected returns. For individual stocks the estimation is conducted in windows, generating a time-series

of component estimates. A bootstrap method would be computationally costly, so I elect to use the t-statistic of the time-series component estimates to measure statistical significance.

4.4 Biased estimation and a bootstrap experiment

Conrad & Kaul (1998) find that variation in the unconditional expected return estimates for individual stocks exceed momentum returns by a large margin. Thus, the time-series components seemingly contribute negatively to momentum returns, while all momentum returns seem to be driven by variation in unconditional expected returns. Jegadeesh and Titman (2002) show that the results in Conrad & Kaul (1998) are affected by what they call a small-sample bias. Adapting Jegadeesh and Titman (2002), when unconditional expected returns for the 11-month cumulative returns and monthly returns are estimated with error, such that:

$$\tilde{\mu}_i = \mu_i + \varepsilon_{\mu i} \text{ and } \tilde{\gamma}_i = \gamma_i + \varepsilon_{\gamma i} , \quad (8)$$

where $\tilde{\mu}_i$ and $\tilde{\gamma}_i$ are the estimated 1-month and 11-month expected returns and $\varepsilon_{\mu i}$ and $\varepsilon_{\gamma i}$ their respective error terms. Indeed, $\tilde{\mu}_i$ and $\tilde{\gamma}_i$ are unbiased estimators of the true unconditional expected returns, since $E(\varepsilon_{\gamma i}) = E(\varepsilon_{\mu i}) = 0$. However, the cross-sectional covariance estimate is biased upwards.⁸ To see this, note that the cross-sectional covariance estimate $\sigma_{\tilde{\mu}\tilde{\gamma}}$ is given by:

$$\sigma_{\tilde{\mu}\tilde{\gamma}} = \sigma_{\mu\gamma} + \sigma_{\varepsilon_{\mu}\varepsilon_{\gamma}} , \quad (9)$$

⁸ Earlier, we presented the cross-sectional variation in expected returns -component of the return as the covariance between the 11-month cumulative return and the monthly return. Jegadeesh and Titman (2002) use holding period and of the same length, which simplifies the equation to being the cross-sectional variance.

where $\sigma_{\mu\gamma}$ is the true covariance between the 11-month and 1-month expected returns and $\sigma_{\varepsilon_{\mu}\varepsilon_{\gamma}}$ is the covariance between the estimation errors. Because the 11-month and 1-month expected return estimates are positively correlated, it follows that $E(\sigma_{\varepsilon_{\mu}\varepsilon_{\gamma}}) > 0$. To see why Jegadeesh and Titman (2002) call this a small sample bias, note that the covariance of the error terms approaches zero as the amount of observations increases, i.e. $\sigma_{\varepsilon_{\mu}\varepsilon_{\gamma}} \rightarrow 0$, when $T \rightarrow \infty$.

To illustrate this small-sample bias, I replicate a bootstrap experiment conducted by Jegadeesh and Titman (2002). First, I randomly sample from each assets return history, forming a new time-series of returns for the asset. This procedure is designed to scramble the returns such that there is no time-series predictability in the return series. I then form momentum strategies that should, by construction, depend only on the cross-sectional variance of the unconditional expected returns of the assets. Repeating the procedure 500 times results in a bootstrap distribution of momentum returns, when returns have no time-series predictability. If momentum is driven by variation in unconditional expected returns, then returns should be high for these momentum strategies as well.

Jegadeesh and Titman (2002) show that in such a setting sampling with replacement causes individual extreme return observations to enter into the return series both in the formation and holding period. These occurrences cause abnormally high returns to momentum in the artificial return series. When sampling without replacement, Jegadeesh and Titman (2002) find that momentum profits in the scrambled return series are very small, and much smaller than true momentum profits. To replicate this analysis, I construct the bootstrap experiment using sampling both with and without replacement.

There is some discussion in statistics literature of the sensitivity of sample autocovariance estimates to outlier observations. For example, Chan (1992) shows that very few additive outliers can cause information in sample autocovariance estimates to decay completely. He also shows that as few as two innovational outliers can cause

bias, that is not fully eliminated even as $T \rightarrow \infty$.⁹ Stock returns are known to exhibit extreme returns fairly frequently, and return histories for individual stocks are fairly short. This suggests that decomposition estimates of auto- and cross-covariances may suffer from biases.

I test for the existence of bias in the estimates of the contributions of auto- and cross-covariances by extending the bootstrap setting described above, by decomposing momentum returns for each bootstrap replication. Because the random sampling removes any true time-series dependence from returns, autocovariances and cross-covariances should have no significant contribution to the momentum returns in the bootstrap. However, if auto- or cross-covariances have significant positive or negative contributions in the scrambled return series, this indicates that auto- and cross-covariance estimates are biased. Here, the results for the bootstrap, sampling without replacement is more informative of the magnitude and direction of these biases. When sampling with replacement, we encounter the previous problem of extreme observations potentially being present in both formation and holding periods.

For the portfolio datasets, I randomly sample return observations for all of the portfolios jointly. This allows for maintaining the contemporaneous covariance structure between the portfolios. For the portfolio datasets, I then construct a full decomposition for each bootstrap replication, reporting the average contributions. Because individual stocks have varying and non-overlapping return histories, I instead randomly sample for each asset separately. Thus, a complete decomposition is impossible. Instead, I compute return autocovariance for each stock separately, and compute the contribution of autocovariance as the average return autocovariance, weighted by the number of observations on each stock. I then report the average contribution of autocovariances from the bootstrap replications. While this does not enable us to compare autocovariances and cross-covariances, it allows us to analyze whether the bias in autocovariances might be present in individual stocks as well.

⁹ For further discussion on sample autocovariance estimates and outliers, see e.g. Chan (1992, 1995) and Deutsch, Richards & Swain (1990).

Statistical significance for all component estimates is measured by t-statistics from the bootstrap replications.

4.4.1 Adjusting for extreme return observations with winsorization

I examine the relationship between the biases and extreme returns, by estimating the original decomposition and the bootstrap experiment again with returns series, where extreme return observations are pulled in. Because the goal here is not to form an investable strategy, but to weed out extreme return observations that affect the sample estimates, I elect to use winsorized returns. My choice is motivated by Welch (2017) who shows that returns winsorized at levels between 10% and 20% produce more persistent estimates of standard deviations and market-betas, when using *daily* returns. While winsorization is an unconventional choice for dealing with returns, it pulls in extreme return observations with little computational complexity and affects fewer return observations than shrinkage estimators, which require more computational effort and shrink all returns towards their mean. Furthermore, the extreme return observations are often related to common rather than idiosyncratic events and these common events occur fairly frequently throughout the sample period. This makes it difficult to reduce the effect of extreme returns by focusing on any subset of stocks or specific time period. I provide an alternative method of adjusting for extreme returns with the strategy-based decomposition, discussed in subsection 4.5.

I winsorize returns at an aggressive 10% level in for individual stocks, as suggested by Welch (2017). Because the portfolio datasets consist of aggregated returns, they suffer less from extreme return observations. I winsorize portfolio returns at a more conservative 1% level. To confirm that momentum exists in these winsorized return series, I recompute momentum strategy returns for these new series. After this, I repeat the decomposition procedure. I also rerun the bootstrap experiment described above to examine the extent to which winsorization alleviates the biases. If the sensitivity of sample estimates to outliers is the cause of the biases, the biases should be less pronounced in the bootstrap experiment and provide a more accurate representation of the drivers of momentum in the decomposition.

4.5 Strategy-based decomposition

My main method of adjusting for extreme return observations is a strategy-based decomposition novel to the literature. I solve the Lo and MacKinlay (1990) decomposition for three individual investment strategies, each corresponding to one element of the decomposition. This provides a simple and computationally feasible solution to this specific problem and gives a new interpretation to each component as an investable strategy.

Take the relative-weight momentum strategy, where assets are weighted by their past 11-month returns (skipping the month prior to the holding period) relative to the return of their equally weighted index over the same period, i.e. $w_{i,t} = \frac{1}{N} (r_{t-2}^{11} - \bar{r}_{t-2}^{11})$. Using the same notation as in subsection 4.2, we can rewrite the strategy weights as:

$$w_{i,t} = \frac{1}{N} \left(\frac{N-1}{N} (r_{t-2}^{11} - \mu_i^{11}) \right) - \frac{1}{N} \left(\bar{r}_{t-2}^{11} - \bar{\mu}^{11} - \frac{1}{N} (r_{t-2}^{11} - \mu_i^{11}) \right) + \frac{1}{N} (\mu_i^{11} - \bar{\mu}^{11}). \quad (10)$$

Now, substituting $w_o \equiv \frac{1}{N} \left(\frac{N-1}{N} (r_{t-2}^{11} - \mu_i^{11}) \right)$, $w_c \equiv -\frac{1}{N} (\bar{r}_{t-2}^{11} - \bar{\mu}^{11} - \frac{1}{N} (r_{t-2}^{11} - \mu_i^{11}))$ and $w_{\mu\gamma} \equiv \frac{1}{N} (\mu_i^{11} - \bar{\mu}^{11})$ in equation (10), we get:

$$w_{i,t} = w_o + w_c + w_{\mu\gamma}. \quad (11)$$

The returns to the momentum strategy can be written as

$$\pi_t = w_{i,t}r_{i,t} = w_o r_{i,t} + w_c r_{i,t} + w_\mu r_{i,t} \quad (12)$$

Here, we have three strategies. Following Lo and MacKinlay (1990), it is straightforward to show that the expected returns to the strategies in equation (12) correspond to the contributions of their respective component, autocovariances, cross-covariances and cross-sectional variance in unconditional expected returns, i.e.,

$$\sigma_{\mu\gamma} \equiv cov(\mu_i^{11} \mu_i),$$

$$O \equiv \frac{N-1}{N^2} tr(\Omega),$$

$$C \equiv \frac{1' \Omega 1 - tr(\Omega)}{N^2},$$

where Ω is once again the cross-covariance matrix between the current 1-month and past 11-month returns.¹⁰ Now, observing the returns to the strategies formed on these weights, allows us to observe, how autocovariances, cross-covariances and unconditional expected returns contribute to momentum returns.

The strategy weights merit some interpretation. The unconditional expected return – strategy, $\sigma_{\mu\gamma}$, is straightforward. We have a zero-investment strategy that weights assets in proportion to their unconditional expected 11-month return, relative to the unconditional expected 11-month return of the equally weighted market index, going long (short) in assets with high (low) unconditional expected returns. The autocovariance strategy, O , weights assets based on their momentum period return, relative to the unconditional expected return on the asset. The constant $\frac{N-1}{N}$ is inconsequential, but included to account for the fact, that the cross-covariance strategy deducts the assets own return from the equally-weighted index return. The strategy

¹⁰ The full derivations for these decompositions and the strategy weights are shown in Appendix 1

takes on a net long (short) position, when the momentum return of the equally-weighted index is above (below) its unconditional expected return.

The cross-covariance strategy, C , weights assets in proportion to the return of the equally-weighted index return relative to its unconditional expectation. The contribution of the assets own return and expected return to the index and the expected return of the index, respectively are deducted from the weights. In other words, the cross-covariance strategy weights depend on the returns of all assets in the momentum period except itself. As N grows larger, the effect of deducting the asset-specific return does not cause much dispersion in the weights. Now, because the holding period returns add up the equally-weighted index return, the strategy essentially simplifies into a contrarian market timing strategy. The strategy invests long (short), when the equally-weighted index momentum return is below (above) its unconditional expectation.

The most important feature of the strategy-based decomposition is that it allows us to directly tackle the biases caused by extreme returns on unconditional expected returns. Rewriting the momentum strategy into three separate strategies gives an interpretation of the biases as a forward-looking bias. For each of the three strategies, an investor needs an *ex-ante* estimate of the unconditional expected return, μ_i^{11} (henceforth referred to as μ without the subscripts for simplicity), to form the strategy weights. When we have extreme return observations that contribute heavily to the individual unconditional expected returns, it would be easy to accumulate high returns by selecting stocks by estimates of their unconditional expected returns, when these high estimates of unconditional expected returns are caused by large future returns. Writing the components as strategies, we see that this is essentially what happens with the unconditional decomposition. Because the unconditional expected return strategy weights depend positively on μ , a forward-looking estimate introduces positive bias to the role of unconditional expected returns, which is what we see both in the unconditional decomposition, as well as the bootstrap experiment.

The weights on the autocovariance and cross-covariance strategies depend on μ as well. Thus, the approach does not only allow us to interpret the strange behavior of unconditional expected returns in the decomposition, but the negative bias in

autocovariances as well. The autocovariance strategy weights negatively depend on μ and thus, the strategy tends to assign larger (smaller) weights to assets with extreme negative (positive) return realizations over their return history. This leads to large negative returns for the strategy in periods, when these extreme returns occur, causing a negative bias to the component. The cross-covariance strategy-weights depend on μ positively, but to a lesser degree, as the asset specific returns are divided by the number of assets N . When N is large, the positive bias to cross-covariances is likely inconsequential. The interpretation of these biases as forward-looking biases allows us to tackle them directly, by using out-of-sample estimates of unconditional expected returns. In removing return observations, we induce more estimation error, but this is a trivial concern compared to the bias, that including holding period returns into the μ estimates induces.

The second attractive quality of this approach is that it makes it easy to estimate the components for individual stocks, because a cross-covariance matrix is not needed. Thus, returns can be estimated using the whole available universe of stocks. Furthermore, looking at the strategies' performance over time, naturally allows us to look for time-variance in the contributions.

A careful reader will note that the portfolios of cross-covariances and autocovariances are not zero-investment arbitrage portfolios. The strategies dependent on autocovariances and cross-covariances take on time-varying short and long positions depending on the formation period returns of each asset relative to their unconditional expected returns. The autocovariance strategy resembles the time-series momentum strategies explored by Moskowitz, Ooi and Pedersen (2012). Goyal and Jegadeesh (2018) point out that the returns to such strategies are not directly comparable to the cross-sectional momentum strategy. They argue that, because “the differences between these strategies are entirely due to the time-series behavior of equal-weighted index returns, we cannot learn anything about the behavior of individual stock returns from the differences in these two strategies”. Indeed, the differences between these strategies arise from the time-series behavior of the equal-weighted index return. However, Goyal and Jegadeesh (2018) fail to mention the fact that the time-series behavior of the equal-weighted index arises from the autocovariances and cross-covariances of the individual stock returns. While Goyal and Jegadeesh (2018) make

a convincing case against using time-series momentum as a return factor, I argue that the strategies I employ can still be informative of the return mechanism behind momentum.

Furthermore, it is easy to see that these strategies do not accumulate returns by loading on market returns. While the autocovariance and cross-covariance strategies are not long/short arbitrage strategies the expected net long/short position of each strategy is 0. To see this, note that in the long run $E[r_{t-2}^{11} - \mu_t^{11}] = E[\bar{r}_{t-2}^{11} - \bar{\mu}^{11}] = 0$. Returns only arise from a market timing element. The positive autocovariance takes on a net long (short) position, when the equal-weighted index returns over the portfolio formation period are above (below) its expected return, while the cross-covariance strategy takes on an offsetting net short (long) position. The cross-covariance strategy depends mostly on the equal-weighted index return, and as such its returns are informative of the contribution of the market timing element. Because the positions offset, the returns to the positive autocovariance strategy are driven by its market timing element only to the extent that these returns show up of the opposite direction in the cross-covariance strategy.

4.5.1 Estimation methods for the strategy-based decomposition

With the strategy-based decomposition, the average contributions of each component can be now viewed as the average returns to the corresponding strategy. I study the returns to investing into these strategies throughout the sample period and report the average returns as annualized percentage returns. I establish statistical significance using t-statistics.

I discuss above in subsection 4.3, how the investment in the relative-weight strategy is scaled by to be 1\$ long and 1\$ short each period. More accurately, the weights are multiplied at each period t by $s_t = 2 * \sum_{i=1}^N \frac{1}{|r_{i,t-2}^{11} - \bar{r}_{t-2}^{11}|}$. This presents some problems with scaling the weights in the strategy-based decomposition. One option would be to scale the strategy weights for all three decomposition strategies by s_t , in which case the component strategy returns add up to momentum returns each period. However, this leads the autocovariance and cross-covariance -strategies to take on up to 10 times

leverage at times. Because the intuition of the strategy-based decomposition is to view the different drivers of momentum as investable strategies, I elect to instead scale the strategies based on their unique attributes, such that their investment returns are somewhat comparable.

The unconditional expected return strategy is the most straightforward one. Because the weights of the unconditional expected return strategy add up to zero, it is a market neutral long-short strategy strategy weights can be scaled to 1\$ long and 1\$ short each period. However, the autocovariance and cross-covariance strategies are not long/short arbitrage strategies. The autocovariance strategy invests some portion long and some short each period. I scale the autocovariance and cross-covariance strategy weights such, that the strategy always invests 1\$ and distributes it to the long and short legs, based on the strategy weights. In this case, the investment can be 1\$ long (short), when all weights have positive (negative) sign. Furthermore, because the strategies are scaled similarly, the net long and short positions offset each other, leading to a combined long/short strategy, where the cross-covariance strategy offsets the net long/short position of the autocovariance strategy. This method of scaling preserves the important market-timing element of the strategies, while scaling investment to a meaningful level. While investment into these strategies is on average smaller than into the momentum and unconditional expected return strategies, it is important to note that the size of the investment does not affect the statistical significance of the returns.

I explore the robustness of my results to alternative scaling schemes in subsection 5.6.2. First, I include not scaling any strategy at all, in which case the components correspond to true cross-covariance and estimates also add up to momentum returns. This introduces large variation in the investment to both momentum and all component strategies. Second, I scale the strategies by the scaling of the momentum strategy, in which case the component strategy returns should again add up to exactly the momentum strategy returns. This introduces large variation in the investment into the component strategies, while investment into momentum remains the same.

Last, I try scaling the autocovariance and cross-covariance strategy weights for each period such, that, for the autocovariance strategy, investment is 1\$ long and less short, when the momentum period returns were on average positive and 1\$ short and less

long, when they were negative. Because the cross-covariance strategy weights are of the same sign every period, this means that the strategy is 1\$ long, when the equally weighted index return momentum is above its historical mean, and 1\$ short when it is below. For the autocovariance strategy, investment ranges from 1\$ long if momentum for all stocks is above the stocks' historical mean momentum returns, to 1\$ short if momentum returns for all stocks are below their respective historical means, to 1\$ long and 1\$ short, when momentum is above historical mean for half of the stocks and below for half of the stocks.

Because the magnitude of component returns is affected by scaling I also run return regressions with momentum as the dependent variable and component strategies as independent variables. Because the strategies have high correlations with each other, regressing momentum returns on all the components jointly, results in a multicollinearity problem, and we cannot make meaningful inference on whether one of the components subsumes the other by regressing momentum returns on all the strategies and looking at the regression coefficients. However, we can test the extent to which the excess returns in the momentum strategy are captured by any one component strategy, by regressing momentum returns on component strategies individually. I report full regression intercepts and coefficients along with respective t-statistics for the individual stocks, but focus only on regression intercepts and t-statistics for the portfolio datasets. The regression intercepts are reported as annualized percentage returns and can be interpreted as the return not explained by an individual component strategy. To measure the statistical significance of the coefficients, I follow common practice in asset pricing literature and compute t-statistics using standard errors corrected for autocorrelation and heteroskedasticity with the robust estimator suggested by Newey and West (1987).

Forming the component strategy weights also requires an ex-ante estimate of the unconditional expected return μ . When using a lot of observations, we can obtain more accurate estimates of the unconditional expected return. However, return histories are short for individual stocks and, in doing so, we eliminate a part of the investable universe of stocks. For the individual stocks, I elect to use an expanding window up to the portfolio holding period. I require 24 individual observations of momentum returns to form initial estimates of the unconditional 11-month expected returns. This sums up

to a total of 36 months of return history required before a stock enters into the strategy. Because there are many stocks with very short return histories in the sample, 25% of the stocks with the shortest return histories are excluded from my sample. Because return histories to the portfolios extend throughout the sample, we can afford a longer estimation period. I elect to use the first 120 months to form initial estimates of μ and update estimates in an expanding window from there on. In subsection 5.5., I investigate, whether the mechanism behind momentum changes over time, by looking at the time-variance of the returns to the component strategies. For this purpose, I use 120-month rolling estimates for the unconditional expected returns, again requiring 24 months to form initial estimates. This allows us to somewhat relax the assumption of mean-stationary returns, that the use of an expanding estimation window imposes.

Later in robustness checks I show, that my results are not sensitive to the choice of return history required, or using rolling estimates of μ . In robustness checks, I also consider return observations from post-holding periods in forming the μ estimate. This makes the strategies not investable in real time, but increases the return observations available, to improve accuracy of μ estimates. Excluding only the holding period returns from the μ estimates gives us the maximum possible amount of return observations, without including the holding period return in the estimate.

5 EMPIRICAL RESULTS

In this section, I present my empirical results. I begin by estimating decompositions for individual stocks and portfolios of stocks in subsection 5.1, using the methods described in subsection 4.3. In subsection 5.2, I post evidence of biased estimation in the decomposition, using the bootstrap experiment described in subsection 4.4. I present decomposition and bootstrap results using winsorized returns in subsection 5.3. In subsection 5.4, I present the main results for the strategy-based decomposition, discussed in subsection 4.5. I explore the potential time-variance in the return mechanism behind momentum in subsection 5.5. Finally, I test the robustness of my results to a variety of methodological choices in subsection 5.6.

5.1 Momentum returns and the unconditional decomposition

Table 2 presents the realized returns and results for the standard decomposition of momentum for the full sample from January 1926 to June 2018. Returns to the relative-weight momentum strategies are highly statistically significant and economically large for all datasets. The returns to momentum in individual stocks are largest, around twice as large as to momentum in the portfolio datasets. For each of the datasets, the expected returns, calculated as the sum of the decomposition components, match the realized returns relatively well. This suggests that scaling the strategy weights as discussed in subsection 4.3 does not drive major differences between the realized returns and an unconditional estimate of the expected return for these strategies.

The decomposition of momentum in individual stocks is difficult to interpret. All individual components appear to contribute positively to momentum returns with similar magnitude and statistical significance. The estimate for the contributions of cross-covariances accounts for around 26.5% of the total expected momentum return. The estimate for the annualized return contribution of autocovariances accounts for 36.1% of momentum returns. These estimates somewhat surprisingly suggest that autocovariances are in general positive, while cross-covariances are in general negative.

Table 2. Momentum strategy returns and unconditional decomposition results, in the sample period from 1926 to 2018, all datasets.

Dataset	Mom	$E[Mom]$	C	O	$\sigma_{\mu\gamma}$
CRSP	11.90** (6.04)	11.04**	2.93**	3.99**	4.14**
FF6	5.16** (4.12)	5.61**	-0.68	4.71*	1.58**
FF25	4.08** (2.78)	2.36**	-2.06**	2.82*	1.60**
Ind10	4.70** (3.16)	4.94**	-1.88**	6.61**	0.21**
Ind30	7.04** (4.47)	6.14**	-0.05	5.91**	0.29**

The returns are reported as annualized percentage returns. Mom = realized momentum returns, $E[Mom]$ = the expected return of momentum, measured as the sum of the components, C = the contribution of cross-covariances to momentum, O = the contribution of autocovariances to momentum, $\sigma_{\mu\gamma}$ = the contribution of cross-sectional variation in unconditional expected returns.

CRSP = Individual US stock data, FF6 = Six market capitalization and book-to-market sorted portfolios, FF25 = 25 market capitalization and book-to-market sorted portfolios, Ind10 = 10 industry portfolios, Ind30 = 30 industry portfolios.

Statistical significance is denoted by stars * = 0.05 level of significance ** = 0.01 level of significance.

The estimate for the contribution of unconditional expected returns is the largest of the three, accounting for 37.5% of the estimated expected returns. The estimate for the cross-sectional variation in unconditional expected returns is much smaller relative to momentum returns than those found by Conrad & Kaul (1998) and Jegadeesh and Titman (2002). This is likely driven by methodological differences. The bootstrap methodology that I employ rules out all stocks with return histories shorter than 36 months. Furthermore, I use an 11-month formation period, and a one-month holding period with a month skipped in between, while Conrad and Kaul (1998) and Jegadeesh and Titman (2002) only consider strategies that have identical and consecutive holding periods and formation periods.

The significant contributions of return cross-covariances and cross-sectional variation in unconditional expected returns are difficult for theories on momentum to reconcile. The very high contribution of unconditional expected returns suggests a risk-based interpretation of momentum. Lewellen (2002) suggests two theories where momentum can be driven by negative cross-covariances. However, theories suggested by Lewellen (2002) assume that return autocovariances would negatively contribute to momentum returns. Here, we observe a positive contribution. The biases in the

decomposition, discussed in subsection 4.4, can somewhat reconcile these puzzling results. I examine these biases empirically in subsection 5.2.

The results in table 2 for the portfolio datasets are more in line with theoretical literature. Positive autocovariances seem to explain most of momentum returns in portfolios. The autocovariance contributions are positive and statistically significant for each dataset. The contributions of autocovariances account for 84% to 134% of momentum returns in the portfolio datasets. Cross-covariances contribute negatively for all portfolio datasets, but cross-covariance estimates differ significantly from zero only for the 25 Fama and French characteristic-sorted portfolios and the 10 industry portfolios.

My results differ from Lewellen (2002), who finds negative cross-covariances and autocovariances using a similar 12-month horizon. The difference between these results and Lewellen (2002) is likely explained by empirical choices. Lewellen (2002) studies momentum strategies that do not skip one month in between the portfolio formation period and the holding period. Furthermore, Lewellen (2002) studies portfolio returns from 1940 to 1999. Chen and Hong (2002) include the early part of the sample period in their results and find results similar to mine, although they use 6-month formation and holding periods. Thus, including the early part of the sample from 1926 to 1940 seems to tilt results in favor of positive autocovariances. This is somewhat counterintuitive, because momentum strategies are not profitable post-2000 or pre-1940. This suggests that autocovariances are positive during periods when momentum is not profitable and vice versa, supporting the role of negative cross-covariances in explaining momentum.

For all portfolio datasets the contributions of unconditional expected returns are positive and statistically significant but small compared to the contribution of return autocovariances. Variation in unconditional expected returns accounts for a significantly larger proportion of momentum returns in the Fama and French portfolios than in the industry portfolios. This is intuitive, since the Fama and French portfolio returns are related to the size- and value-factors and the portfolio sorting is specifically designed to bring out cross-sectional variation in expected returns.

5.2 Bootstrap evidence of biased estimation

The results above for portfolios of stocks seem most supportive of positive autocovariances as the main driver of momentum, consistent with the implications of most popular theories. However, for individual stocks results are ambiguous, as all components appear to be about equally important in explaining momentum returns. This inconsistency might be attributable to biased estimation. Statistical literature documents that outlier observations give rise to biases in sample autocovariance estimates (see e.g. Chan (1992, 1995) and Deutsch, Richards & Swain (1990)), while Jegadeesh and Titman (2002) show that sample estimates of cross-sectional variance in expected returns are biased upwards by extremes returns and short return histories. Furthermore, intuition provided by the strategy-based decomposition in subsection 4.5 suggest that the contribution of unconditional expected returns may be biased upwards, the contribution of autocovariances may be negatively biased to a similar degree and cross-covariances may be biased positively, but to a much smaller degree than the other two components.

I examine whether these biases exist in the decomposition empirically by replicating a bootstrap experiment conducted by Jegadeesh and Titman (2002). The bootstrap experiment is designed to remove any time-series dependence from the return series. Thus, momentum in these return series' can only be driven by cross-sectional variation in unconditional expected returns. I repeat the bootstrap both sampling with and without replacement to illustrate the upward bias to the contribution of unconditional expected returns. In addition, I decompose momentum in these bootstrap replications to examine whether the time-series component estimates suffer from biases. The bootstrap experiment is described in detail in subsection 4.4.

Table 3 presents the average returns from 500 bootstrap replications for the 1926 to 2018 sample period. Panel A presents results using sampling *with* replacement. For the individual stocks in the CRSP data sample, I find results similar to Conrad and Kaul (1998) and Jegadeesh and Titman (2002). The annualized percentage returns to momentum, using scrambled return series, are statistically significant at the 1% level and almost as large as true realized momentum returns over the same sample period. This seemingly suggests that most of momentum returns can be explained by cross-

sectional variation in unconditional expected returns. However, Jegadeesh and Titman (2002) show that this is due to the procedure of sampling with replacement. When we sample with replacement, individual extreme return observations can enter the sample in both portfolio formation and holding periods, causing a positive bias into the returns.

Table 3. Average momentum strategy returns and decomposition estimates from 500 bootstrap replications under the null hypothesis of no time-series dependence in returns, 1926 to 2018 sample period.

Dataset	Mom	C	O	$\sigma_{\mu Y}$
Panel A: Returns sampled with replacement				
CRSP	8.08**	-	-2.59**	-
FF6	1.01**	2.08**	-2.37**	1.79**
FF25	1.14**	1.33*	-1.72**	1.80**
Ind10	0.17**	0.68*	-0.95**	0.52**
Ind30	0.27**	0.03	-0.48	0.73**
Panel B: Returns sampled without replacement				
CRSP	-0.38	-	-2.35**	-
FF6	0.83**	0.87	-1.17*	1.63**
FF25	0.88**	1.90**	-2.26**	1.44**
Ind10	-0.12**	0.82**	-1.13**	0.24**
Ind30	-0.05	0.43	-0.83**	0.33**

Returns and component estimates are reported in terms of annualized percentage returns. Mom = momentum returns, C = the contribution of cross-covariances to momentum, O = the contribution of autocovariances to momentum, $\sigma_{\mu Y}$ = the contribution of cross-sectional variation in unconditional expected returns.

CRSP = Individual US stock data, FF6 = Six market capitalization and book-to-market sorted portfolios, FF25 = 25 market capitalization and book-to-market sorted portfolios, Ind10 = 10 industry portfolios, Ind30 = 30 industry portfolios.

Statistical significance is denoted by stars * = 0.05 level of significance ** = 0.01 level of significance.

Panel B of table 3 presents the same results using sampling *without* replacement. Now, that extreme returns do not contribute to momentum returns in the same way, the returns to momentum in individual stocks are now statistically indistinguishable from zero. This result, consistent with Jegadeesh and Titman (2002), suggests that the cross-sectional variation in unconditional expected returns plays virtually no role in momentum in individual stocks and that estimates of the contribution of unconditional expected returns are positively biased. This finding is also consistent with intuition provided by the strategy-based decomposition.

The portfolio returns consist of aggregated returns and thus do not exhibit such extreme return observations, so results are similar but less dramatic. When sampling

with replacement, momentum returns in the scrambled bootstrap are positive and statistically significant for all portfolio datasets but economically small. When sampling without replacement, momentum returns are positive only for the Fama and French portfolios. This is, again, consistent with the fact that the Fama and French portfolio returns are related to size and value factor returns and produce significant variation in unconditional expected returns. Strikingly, for the 10 industry portfolios, returns are negative and statistically significant at the 1% level. This effect is economically small but puzzling. A potential explanation for such an effect is that industries, which generate large positive (negative) returns at times, have lower (higher) returns on average.¹¹

The other columns on table 3 report average decomposition estimates from the bootstrap replications. I focus on the results when sampling with replacement, because they are more representative of how the auto- and cross-covariance component estimates are affected by extreme returns. Because the individual stocks cannot be sampled in a way that maintains the contemporaneous covariance structure, I only report full decomposition results for the portfolio datasets. For individual stocks, I only report the contributions of autocovariances.

For all datasets, we observe a statistically significant negative contribution by autocovariances, providing strong evidence that sample estimates of autocovariances are negatively biased. Although direct comparison is difficult due to methodological differences, the bootstrap results suggest that biased estimation contributes to the negative autocovariances found by Lewellen (2002), Pan (2010) and Du (2012). The contributions of cross-covariances provide some evidence of a positive bias to cross-covariance contribution estimates. This positive bias in the contribution of cross-covariances is smaller in magnitude than the negative bias to autocovariances for each of the datasets. Furthermore, the contributions are only statistically significant for the

¹¹ Unreported tests support this intuition. For the 10 industry portfolio dataset, average returns have a strong cross-sectional correlation with the smallest quantiles (smaller than 1st percentile) of returns. The correlation is also notably stronger than the correlation between average returns and the largest quantiles of returns. This suggests that occasional large *negative* returns have a closer connection to large average returns than large *positive* return observations do. Furthermore, this relationship is unique to the 10 industry portfolios among the portfolio datasets.

10 industry portfolios and 25 Fama and French portfolios. Results for the time-series components are consistent with the predictions provided by the strategy-based decomposition.

5.3 Decomposition with winsorized returns

Above, I argue that the bias exhibited in the bootstrap is caused by extreme returns. The bias can, of course, potentially arise due to some other reasons. I now empirically examine this potential link between the biases in the decomposition and extreme return observations by decomposing momentum in return series, where extreme return observations are pulled in by winsorizing the returns. The decomposition results should also give a more representative picture of the mechanism driving momentum. However, my main method for unbiased estimation of the decomposition is the strategy-based decomposition, discussed in subsection 5.4.

The return column in Panel A of table 4 reports annualized percentage returns to (uninvestable) momentum strategies formed using these winsorized return series. The returns are winsorized at the 90% level for individual stocks and at the 1% level for the portfolio datasets. The 5% and 95% monthly return breakpoints are -19% and 22.75% respectively for individual stocks. For the portfolio datasets, the monthly return breakpoints end up between -24.6% and -20.0% for the lower bound and between 20.8% and 32.1% for the upper bound. Momentum returns are positive and statistically significant across all datasets. The returns are also slightly larger for each dataset than momentum returns in the true return series. This suggests that the most extreme return observations drive momentum, but instead work against it to some degree.

The component estimates in Panel A of table 4 provide consistent and strong support for the hypothesis that momentum is driven by positive return autocovariances. Across all datasets, autocovariances contribute positively and with high statistical significance. For individual stocks, autocovariances account for around 75% of the returns to momentum, while the contribution of cross-covariances is statistically insignificant. For the portfolio datasets, the results change even more drastically. The contributions of autocovariances far exceed momentum returns for all portfolio

datasets. The large positive contributions of positive autocovariances are offset to a degree by the negative contributions of positive cross-covariances, which are all statistically significant and economically large.

Table 4. Momentum returns and decomposition estimates in winsorized returns 1926-2018.

Dataset	Return	C	O	$\sigma_{\mu\gamma}$
Panel A: Return decomposition.				
CRSP	12.73**	-0.90	9.54**	3.41**
FF6	5.55**	-7.98**	12.47**	1.05**
FF25	5.35**	-7.74**	10.77**	1.34**
Ind10	5.37**	-6.15**	11.19**	0.18**
Ind30	8.05**	-3.27**	10.28**	0.30**
Panel B: Scrambled bootstrap, sampling without replacement.				
CRSP	3.14**	-	-1.11**	-
FF6	0.65**	0.18	-0.40	1.09**
FF25	1.04**	0.50	-0.71	1.30**
Ind10	-0.07	0.60*	-0.85**	0.22**
Ind30	-0.04	0.37	-0.78**	0.37**

Returns and component estimates are reported in terms of annualized percentage returns.

C = the contribution of cross-covariances to momentum, O = the contribution of autocovariances to momentum, $\sigma_{\mu\gamma}$ = the contribution of cross-sectional variation in unconditional expected returns.

CRSP = Individual US stock data, FF6 = Six market capitalization and book-to-market sorted portfolios, FF25 = 25 market capitalization and book-to-market sorted portfolios, Ind10 = 10 industry portfolios, Ind30 = 30 industry portfolios. Returns for the CRSP dataset are winsorized at the 90% level. Returns for the portfolio datasets are winsorized at the 99% level.

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

Unconditional expected returns contribute statistically significantly and positively for all datasets. Variation in unconditional expected returns account for around 26% of momentum returns for individual stocks and a similar proportion of momentum returns in the Fama and French portfolios. For the 10 and 30 industry portfolios, the contribution of unconditional expected returns is tiny compared to momentum returns.

In Panel B of table 4, I report the results from the bootstrap experiment without replacement, repeated using winsorized returns. Momentum returns are now positive and statistically significant for the individual stocks, which is close to the estimate of cross-sectional variation in unconditional expected returns that we find in the decomposition in Panel A of table 4. Momentum returns for the Fama and French portfolios also remain positive, and statistically significant at the 1% level, suggesting that extreme returns do not drive the differences in unconditional expected returns for

these portfolios. For the industry portfolios, bootstrap returns are small and indistinguishable from zero.

The negative biases to autocovariances are still present and statistically significant at the 1% level for both the industry portfolios and individual stocks. However, the bias to autocovariance estimates for the Fama and French portfolios become statistically insignificant. Furthermore, winsorizing returns does reduce the magnitude of the bias in all datasets. The positive bias in the contribution of cross-covariances is statistically significant for only the 10 industry portfolios dataset.

Overall, the results above suggest that extreme returns are at least partly responsible for the biases in decomposition estimates. An alternative interpretation of the results in table 4 is that returns exhibit reversals around extreme returns and winsorizing returns diminishes these reversal effects, shifting return autocovariances towards being more positive. This notion is supported by the fact that returns to momentum are improved in winsorized returns for each dataset, although the increases in returns are quite small. Furthermore, winsorizing returns does not reconcile the biases in the decomposition particularly well. I now move on to the strategy-based decomposition, which provides an alternative method to estimate the decomposition without bias.

5.4 Strategy-based decomposition of momentum returns

The strategy-based decomposition focuses on, whether an investor can utilize the different drivers of momentum by targeting them directly, using ex-ante information. To establish the validity of the strategy-based decomposition framework, I begin by briefly studying the decomposition using full in-sample estimates of unconditional expected returns, μ . Results should be similar to the results from the unconditional decomposition in subsection 5.1. I then re-estimate strategy returns, when unconditional expected returns are estimated using only information available ex-ante. Because the strategy weights here are scaled individually to form realistic investable strategies, the individual components do not add up to the strategy returns, so inference is more reliant on the sign and statistical significance of the components and return regressions, in which momentum returns are regressed on component strategy returns.

5.4.1 Strategy returns using full sample estimates

The strategy returns using full return histories, including the formation period and holding period returns, to form unconditional expected return estimates are presented in table 5. The results on portfolio data are somewhat consistent with the decomposition results presented in subsection 5.1. Returns to the autocovariance strategies are positive for all datasets but are statistically significant only for the two industry portfolio datasets. Returns to the cross-covariance strategy appear negative but are statistically indistinguishable from zero across all datasets.

Table 5. Returns to momentum and component strategy returns, using full-sample estimates of the unconditional expected return from the 1926 to 2018 sample period.

Dataset	Return	C	O	$\sigma_{\mu\gamma}$
CRSP	11.90**	-0.46	3.14	27.12**
	(6.04)	(-0.26)	(1.41)	(24.34)
FF6	5.16**	-0.42	2.11	5.17**
	(4.12)	(-0.20)	(0.92)	(4.18)
FF25	4.08**	-0.39	1.38	6.06**
	(2.78)	(-0.17)	(0.57)	(5.47)
Ind10	4.70**	-1.77	4.03*	2.26
	(3.16)	(-1.05)	(2.15)	(1.88)
Ind30	7.04**	-0.93	3.87*	3.02**
	(4.47)	(-0.54)	(2.00)	(4.29)

Returns to all strategies are reported in terms of annualized percentage returns. Respective t-statistics are reported below the returns in parenthesis.

Return = returns to the momentum strategy, C = returns to the cross-covariance strategy, O = returns to the autocovariance strategy, $\sigma_{\mu\gamma}$ = returns to the unconditional expected return strategy.

CRSP = Individual US stock data, FF6 = Six market capitalization and book-to-market sorted portfolios, FF25 = 25 market capitalization and book-to-market sorted portfolios, Ind10 = 10 industry portfolios, Ind30 = 30 industry portfolios.

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

Returns to the unconditional expected return strategies are statistically significant for all datasets, except for the 10 industry portfolios. Returns to the unconditional expected return -strategy are very large in each dataset compared to what we find using the standard decomposition. For individual stocks, the returns far exceed momentum returns, in line with the findings of Conrad and Kaul (1998).

There are two main reasons as to why the strategy returns differ from the unconditional decomposition results discussed in subsection 5.1. First, for individual stocks, unlike the rolling window methodology used in subsection 5.1, the strategy-based decomposition does not exclude stocks with shorter return histories, so the sample includes stocks with shorter histories. The second reason is the fact that we scale the weights individually. Because the cross-sectional variation in unconditional expected returns is smaller than the cross-sectional variation in individual return observations, the “investment” into a momentum strategy towards unconditional expected returns is smaller than in a strategy that targets unconditional expected returns directly.

5.4.2 Strategy returns using out-of-sample estimates

Table 6 presents annualized percentage returns and respective t-statistics (in parentheses) for momentum and component strategies that are investable in real time, i.e. using an ex-ante estimate of the unconditional expected return. Estimates are computed in an expanding window for all datasets, with 24 observations of momentum required for individual stocks and 120 observations required for the portfolio datasets. The results for individual stocks and industry portfolios support the hypothesis that momentum is driven by positive return autocovariances. Returns to the autocovariance strategy in these datasets are positive, statistically significant and economically large, albeit markedly smaller than returns to momentum. Return autocovariances are also the only component with positive returns. The cross-covariance strategy earns a small negative return, statistically indistinguishable from zero for individual stocks and both industry portfolio datasets. Returns to the unconditional expected return strategy are statistically indistinguishable from zero for the industry portfolio datasets and statistically significantly negative for individual stocks.

The negative returns to the unconditional expected return strategy for individual stocks merit further attention. A zero-investment *contrarian* strategy that is long stocks with low historical returns and short stocks with high historical returns earns an economically large and highly statistically significant return. Of course, cross-sectional variation in true unconditional expected returns by definition cannot contribute negatively to momentum. However, the strategy returns suggest that taking advantage of the cross-sectional variation in unconditional expected returns is not

feasible for an investor in real time. Furthermore, it is unlikely that the past 11-month returns are a better proxy for the unconditional expected returns than full return histories of stocks. Consistent with the bootstrap experiment results, variation in unconditional expected returns is seems to not be an important driver of momentum.

Table 6. Returns to momentum and component strategy returns, using expanding-window estimates of the unconditional expected, 1926 to 2018 sample period.

Dataset	Mom	C	O	$\sigma_{\mu Y}$
CRSP	10.47** (5.31)	-0.44 (-0.25)	6.01** (2.75)	-4.66** (-3.69)
FF6	4.86** (3.95)	1.30 (0.71)	0.27 (0.13)	4.37** (3.95)
FF25	4.59** (3.34)	1.33 (0.70)	-0.11 (-0.05)	4.34** (5.18)
Ind10	4.15** (2.81)	-1.19 (-0.85)	3.32* (2.07)	-0.25 (-0.22)
Ind30	7.20** (4.63)	-0.09 (-0.06)	3.24 (1.92)	-1.21 (-1.58)

Returns to all strategies are reported in terms of annualized percentage returns. Respective t-statistics are reported below the returns in parenthesis. Unconditional expected returns are estimated using ex-ante available information only, in an expanding window. 36 months are required to form initial estimates for the individual stocks. 120 return month are required to form initial estimates for the portfolio data.

Return = returns to the momentum strategy, C = returns to the cross-covariance strategy, O = returns to the autocovariance strategy, $\sigma_{\mu Y}$ = returns to the unconditional expected return strategy.

CRSP = Individual US stock data, FF6 = Six market capitalization and book-to-market sorted portfolios, FF25 = 25 market capitalization and book-to-market sorted portfolios, Ind10 = 10 industry portfolios, Ind30 = 30 industry portfolios.

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

This effect of lifetime historical returns reverting to the mean is likely closely related to the long-term reversal effect documented by DeBondt and Thaler (1985). Since we have a relatively large amount of stocks with short return histories in the sample, the unconditional expected return strategy ends up being somewhat of a proxy for long-term reversal. Furthermore, the inconsistencies between the decomposition and the strategy returns provide further support for the argument that extreme return observations drive most of the apparent variation in unconditional expected returns for individual stocks.

The Fama and French portfolios provide a more ambiguous picture. Results are inconsistent with both results from the other datasets, and the decomposition results for the Fama and French portfolios in subsections 5.1 and 5.3. Returns to the

unconditional expected return strategy in table 6 are positive, statistically significant and very similar in size to the returns of a momentum strategy over the same period. In other words, an investor is about equally well off investing into the Fama and French portfolios based on their full historical returns as they are investing based on their recent returns. Conversely, an investor cannot earn significant returns on an autocovariance strategy, where unconditional expected return estimates are deducted from momentum returns. Returns to the cross-covariance strategy are also statistically insignificant. The results from the strategy-based decomposition suggest, that momentum in these characteristics-sorted portfolios is likely more a manifestation of size and value factor returns, unrelated to the momentum phenomenon in individual stocks and industry portfolios. The return autocovariances found in the decomposition estimates are likely caused by a different kind of time-series dependence in these factor returns. Further examination of these findings is beyond the scope of this thesis.

5.4.3 Return regressions

While the autocovariance strategy is the only ex-ante investable component strategy with positive and statistically significant returns for the individual stocks and industry portfolios, the returns are much smaller than true momentum returns. Because scaling investment into the strategies is difficult, component strategy returns alone are not fully informative of the return mechanism driving momentum. I provide further evidence on the return mechanism by regressing momentum strategy returns on the component strategies. Because component strategy returns have high correlations with each other, multiple regression would run into a collinearity problem. Instead I regress momentum strategy returns against the individual component strategies separately and compare regression coefficients and regression intercepts. The intercepts are reported in terms of annualized percentage returns and can be interpreted as the return of the momentum strategy left unexplained by the component strategy it is regressed on.

Panel A of table 7 reports regression results for momentum in individual stocks. Regression coefficients between the strategies provide further support for the view that momentum is driven by positive return autocovariances. The regression coefficients between momentum and the time-series components are particularly informative. For individual stocks, the autocovariance strategy has a significant positive coefficient

with the momentum strategy, while the cross-covariance strategy has a large negative regression coefficient with momentum. These coefficients are consistent with positive autocovariances driving momentum. When momentum experiences high returns, the autocovariance strategy returns are high and vice versa.

Table 7. Regression coefficients and intercepts for return regressions

Panel A Momentum returns regressed on component strategy returns, individual stock data, 1926-2018				
Dependent	Intercept	C	O	$\sigma_{\mu\gamma}$
Mom	10.52** (5.84)	-0.41** (-5.59)		
Mom	6.28** (4.25)		0.70** (10.47)	
Mom	12.10** (6.19)			0.35* (2.08)
O	5.39* (2.58)			-0.14 (-0.70)
Panel B Regression intercepts, portfolio data 1926-2018				
Independent	FF6	FF25	Ind10	Ind30
C	5.15** (4.53)	4.96** (4.40)	3.90** (2.98)	7.25** (4.82)
O	4.78** (4.42)	4.63** (4.35)	2.31 (1.90)	5.37** (4.00)
$\sigma_{\mu\gamma}$	4.39** (3.94)	4.31** (3.78)	4.13** (3.05)	7.35** (4.87)

Panel A reports regression coefficients from linear regressions, using strategies formed on individual stocks. The strategy used as the dependent variable in the regression is reported on the first column. Regression intercepts are reported in terms of annualized percentage returns. Respective t-statistics are reported below the returns in parenthesis. The t-statistics are computed using standard errors corrected for heteroskedasticity and autocorrelation, following Newey and West (1987). Mom = momentum strategy, C = cross-covariance strategy, O = autocovariance strategy, $\sigma_{\mu\gamma}$ = unconditional expected return strategy.

Panel B reports regression intercepts from linear regressions, using strategies formed on portfolio data. All regressions in panel B have returns to momentum strategy as the dependent variable. The first column reports the strategy used as the independent variable.

FF6 = Six market capitalization and book-to-market sorted portfolios, FF25 = 25 market capitalization and book-to-market sorted portfolios, Ind10 = 10 industry portfolios, Ind30 = 30 industry portfolios. Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

The negative coefficient between the cross-covariance strategy and momentum is explained by the fact, that the auto-covariance and cross-covariance strategies are almost perfectly negatively correlated. When stock returns exhibit occasional reversals, returns are negative for the autocovariance strategy, and positive for the contrarian cross-covariance strategy, and when returns exhibit continuations, returns are positive for the autocovariance strategy and negative for the cross-covariance

strategy. This supports the view that cross-covariances and autocovariances move together, consistent with Boudoukh, Richardson & Whitelaw (1994), who contend that cross-covariances across stocks should, to most extent reflect the relationship between asset-specific autocovariances and contemporaneous covariances.¹² Because the contemporaneous covariances across stocks are positive, return continuations result in both autocovariances and cross-covariances being positive. Conversely, reversals result in both being negative.

The regression intercepts in Panel A of table 7 show that none of the individual component strategies does a particularly good job of capturing momentum returns. Momentum earns economically large and statistically significant positive returns over each component strategy, and returns are affected little by adjusting for exposures on the component strategies. Unexplained returns for the autocovariance strategy are the smallest. The returns of the momentum strategy are virtually unchanged when adjusting for exposures to the cross-covariance strategy. When adjusting for exposure to the unconditional expected return strategy, momentum strategy returns increase. The improved returns are intuitive, because the unconditional expected return strategy returns are negative and the momentum strategy has a significant positive regression coefficient with the unconditional expected returns. The intercept returns here can be interpreted as returns to a momentum strategy that controls for the fact, that stocks full return histories tend to revert to the mean.

The poor performance of the component strategies in explaining momentum returns, even in a regression setting is puzzling. In a regression setting, the scaling is unlikely to affect inference as much. A potential explanation is that, unlike the momentum strategy, the component strategies require an estimate of the unconditional expected return. Because of short return histories, extreme return observations and potentially

¹² More precisely, Boudoukh et. al (1994) suggest, that if past returns to asset i contain no additional information beyond asset j 's past returns, the cross-correlation between the assets can be written as $cor(r_{i,t-1}, r_{j,t}) = cor(r_{j,t-1}, r_{j,t}) \times cor(r_{i,t}, r_{j,t})$. They find that this prediction is very consistent with the first-order cross-correlations found in weekly returns across size-sorted portfolios.

time-varying unconditional expected returns, these estimates are bound to be noisy. Intuitively, introducing noise to the strategy weights should affect returns negatively.

The autocovariance strategy weights depend negatively on the unconditional expected return estimates. Thus, a possible explanation for the profitability of the autocovariance strategy in the individual stocks is that it picks up its returns by negatively loading on the unconditional expected return strategy. The last row in Panel A of table 7 addresses this concern. Regressing the returns of the autocovariance strategy on the returns of the unconditional expected return strategy results in an insignificant regression coefficient, and a large and statistically significant intercept for the autocovariance strategy. Thus, the positive returns to the autocovariance strategy are not driven by the unconditional expected return strategy.

Panel B of table 7 reports regression intercepts for the portfolio data from similar return regressions. Regression coefficients are left unreported but provide similar inference as above. The component strategies capture even less of momentum returns in portfolios than in individual stocks. When adjusting for exposures to the autocovariance strategy, momentum returns in the industry portfolios are significantly diminished from the returns seen in table 6. Returns are virtually unaffected by adjustments for exposures on the cross-covariance and unconditional expected returns strategies, providing some support for momentum driven by positive autocovariances. For the Fama and French portfolios regression intercepts do not differ much from the original momentum returns, when controlling for any of the component strategies. Intercepts for the unconditional expected return strategy are the smallest. However, the differences in intercepts are too small for reliable inference to be made.

5.5 Time-variance in the return mechanism

The empirical results discussed above do not require that mechanism is consistent over time and the results can be interpreted as the average contributions of component strategies to momentum over the entire sample period. However, the return mechanism may change over time. I now examine this potential time-evolution in the return mechanism, by examining the decomposition over different time periods in the sample. I focus on individual stocks and the strategy-based decomposition, which naturally

allows for time-variance in autocovariances, cross-covariances and unconditional expected returns. I study both strategy returns, as well as return regressions over different time periods.

Figure 1 plots 10-year rolling annualized returns of the component strategies using 10-year rolling estimates of the unconditional expected return. Figure 2 plots the log cumulative returns of the component strategies, using 10-year rolling estimates of the unconditional expected return, as well as momentum returns over the same time-period. Both figures show the component strategy returns are highly volatile pre-1940. Thereafter, the autocovariance strategy yields high and positive returns fairly consistently throughout the entire sample period up until around 1990s. Returns to the autocovariance strategy then begin a steady decline around the time that the effect is initially documented in the literature by Jegadeesh and Titman (1993). Returns to the cross-covariance strategy are mostly negative during the early part of the sample but are mostly positive post-1980. Returns to the cross-covariance strategy decline similarly to the autocovariance strategy post-2000.

The slow decline in time-series predictability of returns might be indicative of increasing arbitrage activity related to momentum. This result is consistent McLean and Pontiff (2016), who find that anomalies tend to attenuate after they are discovered. The results are also consistent with recent evidence by Baltussen, Bakkum and Da (2019), who find that daily index level return autocovariance has declined around the world.¹³

¹³ Baltussen, Bakkum and Da (2019) connect their findings with the increase in index investing. Furthermore, they study daily autocorrelation of the market index, while I focus on a much longer horizon.

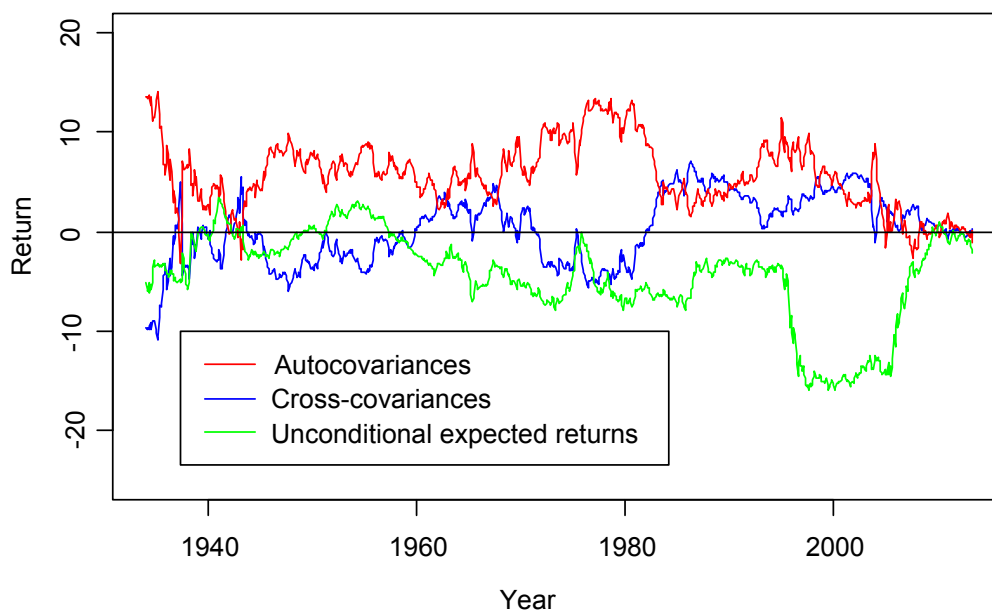


Figure 1. 10-year rolling annualized percentage returns of component strategies.

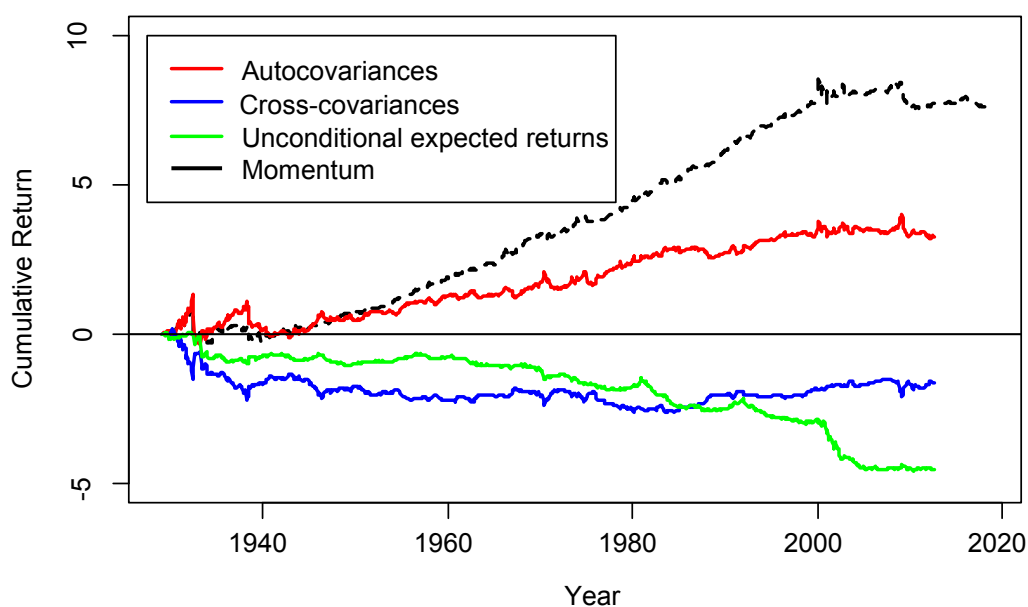


Figure 2. Cumulative returns of momentum and component strategies.

We see from figure 1 that the negative returns to the unconditional expected return strategy are fairly consistent throughout the sample period. However, a significant portion of the negative returns are driven by very poor performance around the turn of the century. The cumulative return plot in figure 2 reveals that these negative returns are related to the stock market crash just after 2000. This is likely related to the tech bubble around the turn of the century. In the late 1990s, many new stocks enter into

the sample. Because stock returns during the late 1990s were very high, the unconditional expected return strategy assigns large positive weights to new stocks, because of their high lifetime returns. During the stock market crash, many of these stocks realize large losses and a significant portion of the companies fail. We can see from figure 2, that the losses from this brief period account for around a third of the cumulative losses of the unconditional expected return strategy over the entire sample period from 1926 to 2018.

Panel A of table 8 presents strategy returns over five 20-year subperiods in the sample. Results largely support the inference above. Autocovariance-strategy returns are positive and statistically significant in only 2 of the 20-year periods, in 1940 to 1959 and 1980 to 1999. These results are consistent with the weaker performance of the autocovariance strategy that we observe in figure 1 during the early part of the sample, in the 1960s and post-2000. Returns to the cross-covariance strategy increase over each 20-year period. The returns are, however, statistically insignificant in all periods, except from 1980 to 1999. Furthermore, the autocovariance strategy returns are slightly larger over the 1980 to 1999 period. The returns to the unconditional expected return strategy are negative in all periods but the negative returns are statistically significant only in the last two 20-year periods from 1980 to 2018.

Overall, the autocovariance and cross-covariance strategy returns suggest that momentum is consistently driven by positive autocovariances. The 1980s, however, seems to mark a notable shift in return cross-covariances, as the returns to the cross-covariance strategy rapidly turn towards positive and also remain mostly positive after this period. This shift is clearly visible in figure 1 as well. At the same time returns to the autocovariance strategy returns decline but remain positive. It is possible that the return mechanism driving momentum changes. Since both component strategy returns remain positive, intuitively, momentum returns should remain positive as well.

To provide a simple test, of whether cross-covariances explain momentum in the later part of the sample I repeat the return regressions on the early and late parts of the sample separately. If momentum is indeed driven by cross-covariances in the late sample, this should show up in the return regression coefficients. Figures 1 and 2 provide little guidance in selecting a breakpoint, so I elect to split the sample from the

year that Jegadeesh and Titman (1993) initially document momentum. If a structural shift would occur in the mechanism, it could potentially explain the underwhelming regressions results seen in subsection 5.4.3 as well, because a linear regression fails to capture structural changes.

Table 8. Decomposition of momentum returns in different sample periods, momentum in individual stocks.

Sample period	Mom	C	O	$\sigma_{\mu\gamma}$
Panel A: Strategy-based decomposition across different sample periods				
1926-1939	4.48 (0.47)	-8.11 (-0.72)	11.57 (0.92)	-4.41 (-0.71)
1940-1959	10.92** (4.62)	-2.95 (-1.18)	7.37* (2.48)	-0.91 (-0.73)
1960-1979	13.80** (4.26)	-0.31 (-0.10)	7.03 (1.79)	-3.51 (-1.67)
1980-1999	19.86** (6.66)	3.95* (2.17)	5.97* (2.21)	-5.62* (-2.55)
2000-2018	-0.24 (-0.05)	1.90 (0.71)	0.23 (0.05)	-9.05** (-2.70)
Panel B: Strategy return regressions before and after 1993				
1926-1992	10.43** (6.69)	-0.34** (-5.45)		
	7.76** (5.95)		0.65** (14.43)	
	2.77 (0.67)	-0.43** (-4.20)		
1993-2018	1.32 (0.50)		1.14** (8.81)	

Returns to all strategies and component estimates are reported in terms of annualized percentage returns. Respective t-statistics are reported below the returns in parenthesis. In panels A and C, the strategies are formed using unconditional expected returns estimated using ex-ante available information only. The estimates are formed in a 120-month rolling window with 36 return months required to form initial estimates. The sample period column reports the years of the sample period under study.

In panel A: Mom = momentum strategy returns over the period, C = the contribution of cross-covariances to momentum, O = the contribution of autocovariances to momentum,

$\sigma_{\mu\gamma}$ = the contribution of cross-sectional variation in unconditional expected returns.

In panel B: Mom = regression intercept of momentum strategy over component strategy, C = regression coefficient of the cross-covariance strategy and C = regression coefficient of the autocovariance strategy. The intercepts are reported in terms of annualized percentage returns. Respective t-statistics are reported below the coefficients in parenthesis and computed using standard errors corrected for heteroskedasticity and autocorrelation, following Newey and West (1987).

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

Regression results are reported in Panel B of table 8. The momentum column represents regression intercepts, while regression coefficients of the cross-covariance

and autocovariance strategies are reported in their respective columns. Although the cross-covariance strategy is more profitable in the later part of the sample, the regression coefficients are negative and highly statistically significant over both subsamples. Conversely, autocovariance strategy coefficients remain positive and highly statistically significant. We do not observe any structural shift that would be consistent with return cross-covariances explaining momentum returns. To the extent that return cross-covariances turn negative and contribute positively to momentum, this positive contribution is subsumed by autocovariances that also turn negative. In unreported tests, I try a variety of other cutoff points from 1980 onwards for splitting the sample and find similar results. Interestingly, regression intercepts of momentum over both the cross-covariance and autocovariance strategies are economically small and statistically insignificant in the 1993 to 2018 sample period. This is, however, most likely only reflective of the fact that momentum returns are small in the latest part of the sample.¹⁴

5.6 Robustness checks

I conclude the empirical part of my thesis by exploring the sensitivity of my results to a variety of methodological choices. For these robustness checks, I focus on the main dataset consisting of individual stock returns. Following the advice of Fama and French (2008), I first examine my results across different market capitalization groups. For this robustness check, I study both the unconditional decomposition with winsorized returns and the strategy-based decomposition. The rest of the robustness checks focus on methodological choices in constructing the component strategies in the strategy-based decomposition.

5.6.1 Decomposition across market capitalizations

In my main results, I exclude microcap stocks (stocks below the 20th percentile of market capitalization). Table 9 contains decomposition results using the full sample

¹⁴ Indeed, splitting the sample in 1980 instead results in positive and statistically significant regression intercepts.

with these microcaps included, as well as separately across four different market capitalization groups. Panel A of table 9 reports the results of the strategy-based decomposition. Panel B of table 9 reports the unconditional decomposition results using winsorized returns. The results in both panels confirm that my main results are not driven by any particular market capitalization group. For all size groups, except for microcaps, the results from both decompositions are consistent with earlier results. Consistent with prior literature, returns to momentum decrease as we move from smaller to larger market capitalization groups. The estimated contribution of positive autocovariances diminishes in a similar fashion. Returns to the autocovariance strategy are statistically insignificant among the largest market capitalization quintile of stocks.

Table 9. Decomposition of momentum across market capitalizations.

Size	Mom	C	O	$\sigma_{\mu Y}$
Panel A: Strategy-based decomposition				
All stocks	4.64* (2.08)	-0.53 (-0.28)	4.65* (2.00)	-10.08** (-5.96)
Mega	6.23** (3.06)	0.59 (0.38)	2.96 (1.50)	-1.97 (-1.30)
Large	8.50** (4.19)	0.16 (0.10)	4.48* (2.16)	-3.58** (-2.65)
Small	12.93** (6.59)	-1.59 (-0.76)	7.91** (3.24)	-4.74** (-4.19)
Micro	-8.15* (-2.39)	-0.03 (-0.01)	-0.39 (-0.12)	-13.58** (-4.53)
Panel B: Lo and MacKinlay (1990) decomposition, using winsorized returns				
All stocks		-1.23*	7.80**	5.37**
Mega		0.72	2.77**	1.05**
Large		0.58	5.85**	1.97**
Small		-3.78**	12.86**	4.12**
Micro		-6.80**	6.19**	8.80**

Returns to all strategies and component estimates are reported in terms of annualized percentage returns. In Panel A, respective t-statistics are reported below the returns in parenthesis. The strategies in Panel A are formed using unconditional expected returns estimated using ex-ante available information only. The estimates are formed in an expanding window with 36 return months required to form initial estimates.

The size column reports the market capitalization group. All stocks = full sample of individual stocks from the CRSP dataset, with microcaps included, Mega = Stocks in the top 20th percentile of market capitalization, Large = Stocks with market capitalization above the median market capitalization, Small = stocks between the 20th and 50th percentile of market capitalization, Micro = stocks in the smallest 20th percentile of market capitalization.

Mom = momentum returns, C = the contribution of cross-covariances to momentum, O = the contribution of autocovariances to momentum, $\sigma_{\mu Y}$ = the contribution of cross-sectional variation in unconditional expected returns.

In panel B, returns for the CRSP dataset are winsorized at the 90% threshold.

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

The reversals found in the unconditional expected return strategy seem to be more pronounced in smaller market capitalization quintiles and especially large for the microcap stocks. Conversely, the estimates of cross-sectional variation in unconditional expected returns in Panel B of table 9 are positive but much larger for smaller market capitalization stocks. This suggests that much of the variation observed in unconditional expected returns are driven by the smallest stocks in the sample. Furthermore, because the reversals seen in the strategy-based decomposition are stronger, when variation in unconditional expected returns is stronger, it is likely that the unconditional expected return estimates are driven by extreme return observations and short samples. This is likely reflective of the high positive correlation between market capitalization and the length of return history. Smaller stocks have shorter return histories and are more affected by the small-sample bias.

In the microcap quintile, instead of momentum we observe large and statistically significant reversals in returns. The auto- and cross-covariance strategy returns are both essentially zero. The negative returns to the unconditional expected return strategy are larger than momentum returns and highly statistically significant. However, the unconditional decomposition results in Panel B of table 9 suggest the opposite. When we pull in the extreme return observations, we find positive auto- and cross-covariances in returns. The results suggest that microcap stocks have such extreme individual return observations and short return histories, that estimating time-series dependence or unconditional expected returns is very difficult. The unconditional expected return estimates used to form the time-series strategies are not informative of true unconditional expected returns and time-series component strategies trade mostly based on noise.

5.6.2 Robustness to scaling the strategy weights

In subsection 4.5.2., I discuss, how scaling investment in the component strategies, can affect returns. Scaling the investment in the strategies affects the decomposition results, because it imposes conditionality on the cross-sectional dispersion of the weights. Furthermore, the time-series component strategies are not zero net investment strategies and scaling the weights represents a challenge. As a robustness check, I explore alternative methods of scaling investment into the strategies.

Table 10 contains decomposition results using three different methods of scaling the strategy weights. No scaling corresponds to returns to strategies in which strategy weights are not scaled from period to period. The investment size varies based on the return dispersion and the components should correspond exactly to the autocovariances, cross-covariances and unconditional expected returns, with the distinction, that the unconditional expected return is computed using out-of-sample data. Indeed, we can see that the component strategy returns add up to the realized return of the momentum strategy. Returns to momentum are much smaller when investment is not scaled. Returns to an unscaled momentum strategy are very small compared to the return of the scaled strategy. Momentum returns are, however, still statistically significant at the 1% level.

Table 10. Returns to component strategies using different methods of scaling the investment weights in the strategies, from 1926 to 2018.

Scaling method	Mom	C	O	$\sigma_{\mu\gamma}$
No scaling	1.43** (3.38)	-0.30 (-0.28)	2.08 (1.66)	-0.35** (-3.72)
Momentum scaling	10.47** (5.31)	5.48 (0.60)	6.74 (0.65)	-1.75** (-2.86)
Always 1\$ long or 1\$ short	10.47** (5.31)	0.12 (0.05)	9.13** (3.61)	-4.66** (-3.69)

Returns to all strategies are reported in terms of annualized percentage returns. Respective t-statistics are reported below the returns in parenthesis. The strategies are formed using unconditional expected returns estimated using ex-ante available information only. The estimates are formed in an expanding window with 36 return months required to form initial estimates. In the scaling method column, No scaling corresponds to strategies, in which investment is not scaled from period to period, Momentum scaling corresponds to strategies, where investment into the component strategies is scaled based on the scaling of the momentum strategy, by $s_t = 2 * \sum_{i=1}^N \frac{1}{|r_{i,t-2}^{11} - \bar{r}_{t-2}^{11}|}$, and Always 1\$ long or 1\$short corresponds to strategies, where the component strategies are scaled such that they are either \$1 long or \$1 short every period, and the maximum investment is \$1 long, \$1 short.

Mom = returns to the momentum strategy, C = returns to the cross-covariance strategy, O = returns to the autocovariance strategy, $\sigma_{\mu\gamma}$ = returns to the unconditional expected return strategy.

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

Not scaling the strategy weights results in both of the time-series component strategies in taking up to 2.37 times leverage. At other times, when cross-sectional return dispersion is small, there is essentially no investment at all. The component strategy returns without scaled investment do not contradict the main results. Only the returns to the autocovariance strategy are positive. While the autocovariance strategy returns are only statistically significant at the 10% level, the cross-covariance strategy returns

are small by comparison and statistically indistinguishable from zero. The negative returns to the unconditional expected return strategy are much smaller when the investment is not scaled, but remain statistically significant at the 1% level.

Momentum scaling corresponds to strategies in which component strategies are scaled similarly to the momentum strategy, by $s_t = 2 * \sum_{i=1}^N \frac{1}{|r_{i,t-2}^{11} - \bar{r}_{t-2}^{11}|}$. Once again the component strategy returns add up to the momentum returns. Returns to the autocovariance and cross-covariance strategies are large, but statistically indistinguishable from zero. The time-series component strategies now take on positions that are levered up to 9 times. Because the strategy weights are mostly affected by s_t , and the weight dispersion in the cross-covariance strategy is small, the cross-covariance strategy loses its interpretation as a contrarian market-timing strategy and resembles more a market-timing strategy based on cross-sectional return dispersion. Consistent with earlier results, returns to the unconditional expected return strategy are negative and statistically significant, albeit small.

The last row in table 10 alters the scaling to the cross-covariance and autocovariance strategies are scaled such that they are either \$1 long or \$1 short every period, and the maximum investment is \$1 long, \$1 short. For the cross-covariance strategy, this essentially means that the strategy is \$1 short (long), whenever the 11-month cumulative returns of the equally weighted returns are above (below) the mean of historical 11-month cumulative returns of the equally weighted index with no middle ground. Returns to the autocovariance strategy are now highly statistically significant and comparable in size to momentum returns over the same period. Cross-covariance strategy returns are once again indistinguishable from zero. The scaling, and thus, the returns of both momentum and the unconditional expected return strategy are the same as in the main results.

To summarize the results from table 10, the method of scaling component strategy weights is quite influential on the investment into the strategies, and thus the strategy returns. However, the results are most supportive of positive return autocovariances driving momentum, regardless of how investment into the strategies is scaled.

5.6.3 Robustness to methods of estimating the unconditional expected return

Forming the strategy weights requires an estimate of unconditional expected return. In this subsection, I show that my results are not sensitive to any particular choice of estimation window for the unconditional expected return. Table 11 presents the strategy returns and respective t-statistics using different estimates of the 11-month unconditional expected returns. All OOS corresponds to an unconditional expected return estimate of the mean of all 11-month mean cumulative return observations, except for the 11 months, when the holding period returns enter the 11-month cumulative return. This means that the strategies are of course not investable, because they include future information in to the unconditional expected return estimates. Forward-looking corresponds to estimates, where only post-holding return observations are used to form unconditional expected return estimates. 120/24 corresponds to 120-month rolling estimates with an initial 24 months required. 60 required corresponds to investable strategies, where we require 60 observations to form initial estimates on the unconditional 11-month expected return. 12 required corresponds to investable strategies where only 12 observations of the 11-month cumulative return are required.

Momentum returns in table 11 differ from earlier, because the return observations in the beginning of each return series are not consumed to form initial estimates. All momentum returns are, however, similar in magnitude and significance as earlier. The autocovariance strategy earns positive and statistically significant returns at the 5% level across all estimates of the unconditional expected return, except for when we require longer than 60 months of return history. The strategy returns, excluding the 60 month required estimates, are also similar in magnitude regardless of the unconditional expected return estimate used. Returns to the respective cross-covariance strategies are small by comparison and are statistically insignificant in all cases.

Requiring a 60-month initial return history to estimate unconditional expected returns appears to have an effect on auto- and cross-covariance strategy returns. Because return histories are on average short, the 60 month requirement restricts, restricts the number of stocks in the sample to around half of the original. Positive returns to the autocovariance strategy in this sample of stocks are small and only statistically

significant at the 10% level. Cross-covariance strategy returns are positive unlike for all other specifications, but remain statistically insignificant. It appears, that stocks with longer return histories have on average lower autocovariances and negative cross-covariances. Momentum returns are significantly diminished as well, suggesting that momentum is stronger among stocks with shorter return histories. This is likely due to the positive correlation between market capitalization and short return histories and the fact that momentum is stronger in smaller capitalization stocks.

Table 11. Momentum and component strategy returns in individual stocks from 1926 to 2018, using different estimates of the unconditional expected return to form the strategy weights.

Estimate of μ	Mom	C	O	$\sigma_{\mu\gamma}$
All OOS	9.65 (4.76)	-0.54 (-0.29)	5.85 (2.52)	-8.35 (-8.24)
Forward-looking	11.05 (5.46)	-1.82 (-0.99)	7.47 (3.25)	-4.95 (-5.20)
120/24	10.47 (5.31)	-0.50 (-0.28)	6.08 (2.79)	-4.45 (-3.49)
60 required	8.48 (4.36)	1.01 (0.58)	3.59 (1.69)	-4.03 (-3.36)
12 required	11.29 (5.75)	-0.71 (-0.41)	6.67 (3.11)	-3.94 (-3.24)

Returns to all strategies are reported in terms of annualized percentage returns. Respective t-statistics are reported below the returns in parenthesis. The rows represent different methods of estimating the unconditional expected return. All OOS = Unconditional expected return estimates using all return observations, except for the holding period return, Forward-looking = unconditional expected return estimates formed on post-holding period returns, 120/24 = unconditional expected returns estimates using ex-ante available information only, estimated in a 120-month rolling window, with 24 months required to form an initial estimate, 60 required = expanding window, ex-ante estimates of the unconditional expected return with 60 months required to form an initial estimate, 12 required = expanding window, ex-ante estimates of the unconditional expected return with 12 months required to form an initial estimate.

Mom = returns to the momentum strategy, C = returns to the cross-covariance strategy, O = returns to the autocovariance strategy, $\sigma_{\mu\gamma}$ = returns to the unconditional expected return strategy.

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

The unconditional expected return strategy generates highly statistically significant and negative returns across all unconditional expected return estimates. The negative returns are in fact larger, when we include future returns to the unconditional expected return estimates, even when we only use future information. This result is interesting, as it suggests that stocks with positive (negative) individual extreme return observations earn below (above) average return throughout their existence, and returns to the unconditional expected return strategy are not entirely driven by long-term

reversals. This evidence is consistent with findings by Bali, Cakici and Whitelaw (2011) and Boyer, Mitton and Vorkink (2010), who show that investors prefer stocks with positively skewed returns and require a return premium for holding stocks with negatively skewed returns. Further examination of these findings is left for future research.

This effect can also potentially explain the negative bias present in estimates of return autocovariances. For example, consider a stock that realizes a high positive return in one period. This high return realization then results in a higher average return over the stock's lifetime. Now, because the stock's returns in other periods are on average low, the subsequent return is likely to be below the stock's average return. When both time-series predictability and the dispersion in the stocks' returns is smaller in other periods, these single occurrences can drive the sample autocovariance estimates towards being negative.

5.6.4 Other formation and holding periods of momentum

Prior literature on momentum varies in their choice of formation period and holding period for momentum. I conclude the robustness checks by exploring the robustness of my results to these methodological choices. Table 12 presents the strategy-based and unconditional decomposition results, using different variants of momentum. 6/6 represent momentum strategies, where the holding period and formation periods are both six months, and a month is skipped between the formation and holding periods. 6/6 without skipping a month is the same strategy, with adjacent formation and holding periods. 6/1 is a momentum strategy with a six month formation period and a one month holding period, with one month skipped in between. 12/1 without skipping corresponds to a momentum strategy without one month skipped in between formation and holding periods. All returns to momentum strategies are positive, statistically significant and economically large. All returns are, however considerably smaller at around half of the return to the 11-month momentum strategy that skips a month before the holding period.

Table 12. Momentum and component strategy returns in individual stocks from 1926 to 2018 with different holding periods and formation periods used to form the strategy weights.

Strategy	Mom	C	O	$\sigma_{\mu\gamma}$
6/6	5.35 (2.42)	0.30 (0.21)	2.65 (1.40)	-4.72 (-4.50)
6/6 Without skipping a month	4.60 (2.13)	1.77 (1.15)	0.80 (0.39)	-3.74 (-3.20)
6/1	6.39 (3.29)	1.71 (0.97)	1.90 (0.85)	-3.39 (-2.60)
12/1 without skipping	5.83 (2.76)	-1.52 (-0.83)	5.30 (2.36)	-4.35 (-3.45)

Returns to all strategies are reported in terms of annualized percentage returns. Respective t-statistics are reported below the returns in parenthesis. The strategies are formed using unconditional expected returns estimated using ex-ante available information only. The estimates are formed in an expanding window with 36 return months required to form initial estimates. The rows represent momentum and component strategies with different holding and formation periods. 6/6 = strategies using the past 6-month cumulative return as the measure of momentum and a 6-month holding period, with one month skipped between the formation and holding periods, 6/6 Without skipping a month = strategies using the past 6-month cumulative return as the measure of momentum and a 6-month holding period, without one month skipped between the formation and holding periods, 6/1 = strategies using the past 6-month cumulative return as the measure of momentum and a one month holding period, with one month skipped between the formation and holding periods, 12/1 without skipping = strategies using the past 12-month cumulative return as the measure of momentum and a one month holding period, without one month skipped between the formation and holding periods.

Mom = returns to the momentum strategy, C = returns to the cross-covariance strategy, O = returns to the autocovariance strategy, $\sigma_{\mu\gamma}$ = returns to the unconditional expected return strategy.

Statistical significance is denoted by stars: * = 0.05 level of significance, ** = 0.01 level of significance.

While the results in table 12 do not contradict earlier results, the autocovariance strategy returns appear quite sensitive to the choice of formation period. Autocovariance strategy returns are only statistically significant for the momentum strategy with a 12-month formation period, without a month skipped. Returns are rendered insignificant when the return horizon is reduced to 6 months. A likely explanation for the reduced performance of both the autocovariance strategy and momentum at shorter horizons is a spillover effect from short-term reversals. Novy-Marx (2012) finds that, in US individual stocks, momentum returns are mostly driven by the months between 7 and 12 months prior to the holding period. Goyal and Wahal (2015) show that this effect is caused by reversal effects from returns two months prior to the holding period. This effect is stronger for the autocovariance strategy than for momentum, in part due to the autocovariance strategy returns being smaller than momentum returns. Cross-covariance strategy returns remain statistically indistinguishable from zero across all strategy specifications. Returns to the

unconditional expected return strategy are negative and highly statistically significant across all strategy specifications, consistent with earlier results.

6 SUMMARY

In this thesis, I study momentum, one of the most widely documented phenomena in the cross-section of stock returns. I begin with a review of prior literature on momentum in general. I cover empirical evidence on momentum, theories proposed to explain momentum and literature studying cross-sectional and time-series determinants of momentum returns. I conclude my literature review with a discussion on the literature focused on return dynamics that drive momentum. More specifically, the returns to a momentum strategy can be decomposed into three drivers of return: positive return autocovariances, negative return cross-covariances and the cross-sectional variation in unconditional expected return. Understanding which of these components drives momentum returns is important in distinguishing between theories on momentum, because all theories on momentum need to make an implicit assumption on the return mechanism. I review both the theoretical framework and prior empirical evidence on the decomposition of momentum returns.

Empirically, I examine the decomposition of momentum returns in US individual stocks and portfolios of US stocks in a sample extending from 1926 to 2018. I add to prior literature on decomposing momentum returns by suggesting two novel methods that allow for estimating the decomposition for individual stocks. The first one involves estimating cross-covariance matrices and the decomposition in a rolling window. The second is a strategy-based decomposition, based on analytically solving for three separate strategies that depend only on autocovariances, cross-covariances and unconditional expected returns, respectively. This strategy-based decomposition allows me to decompose momentum returns without estimating a cross-covariance matrix.

I implement a bootstrap experiment that provides evidence that the sample estimates of cross-covariance matrices used in prior research suffer from biased estimation. I find that estimates of the contributions of unconditional expected returns and return cross-covariances are positively biased, while estimates of the contribution of return autocovariances are negatively biased. My findings on cross-covariances and autocovariances are novel to the literature. I link these biases to extreme return observations, by examining the decomposition using winsorized returns. I show that

the biases are reduced, when unconditional expected returns and cross-covariances are estimated using winsorized returns. Decomposition estimates of the contributions using winsorized returns suggest that momentum is mostly driven by positive autocovariances in returns.

The strategy-based decomposition provides a different interpretation to these biases as forward-looking biases. Employing the component strategies using ex-ante available information results in an unbiased decomposition. Using the strategy-based decomposition, I find consistent evidence that momentum is driven by positive autocovariances in both individual stocks, as well as portfolios of stocks sorted by industry. I show that these results are consistent over time and robust to a variety of methodological choices.

My results are consistent with behavioral explanations for momentum, where momentum is caused by underreaction to new information by market participants. However, my results do not preclude a rational explanation for momentum. Many rational theories proposed in the literature likewise imply that momentum is driven by positive return autocovariances. While the return mechanism behind momentum is important to assess, further work is needed to distinguish between competing theories on momentum. Different theories generate different testable implications and testing many theories jointly remains an important challenge in better understanding momentum.

While I find consistent evidence across individual stocks and industry portfolios, evidence on what drives momentum in the Fama and French characteristic-sorted portfolios are inconclusive. Evidence from the standard decomposition suggests that momentum in these factor portfolios is driven by positive return autocovariances, but the strategy-based decomposition show that an investor is equally well off investing into the Fama and French portfolios based on unconditional expected return estimates instead. It is likely that, while momentum in individual stocks and industry momentum are similar phenomena, momentum in characteristic-sorted portfolios is somewhat distinct from other forms of momentum. Research dedicated to studying momentum in portfolios related to factor returns is sparse and represents an interesting area to further pursue.

The strategy-based decomposition also allows for time-varying contributions of the return mechanisms to momentum. In analyzing the time-variance I find no consistent evidence of a time-shift in the mechanism behind momentum. I find that, while the contribution of autocovariances declines and the contribution of cross-covariances increases in the latter part of the sample, at the same time, the profitability of momentum declines as well. Furthermore, momentum returns remain positively correlated with the autocovariance strategy. These results suggest that despite negative cross-covariances contributing positively, their effect is subsumed by a negative contribution from negative autocovariances. I also find a steady decline in time-series predictability of returns after 1990s that can explain the poor recent performance of momentum strategies. However, my examination of the attenuated momentum returns and time-series predictability is parsimonious. The recent performance of momentum and its implications for future returns is an important topic for future literature to assess.

I also document an interesting phenomenon, where a stock's lifetime returns *negatively* predict its current returns. A similar reversal effect is present and significant when using pre-holding period returns, post-holding period returns or both in estimating lifetime returns. These findings potentially add to prior literature on long-term reversals in returns, as well as literature on investors' preference for assets with positively skewed returns over assets with negatively skewed returns. Exploring these reversal-effects in lifetime returns are beyond the scope of this thesis. However, further work linking these findings to prior literature is warranted.

REFERENCES

- Ahn, D. -H., Conrad, J. & Dittmar, R. (2003). Risk adjustment and trading strategies. *The Review of Financial Studies* 16(2), 459-485.
- Antoniou, C., Doukas, J. A. & Subrahmanyam, A. (2013). Cognitive dissonance, sentiment, and momentum. *Journal of Financial and Quantitative Analysis* 48(1), 245-275.
- Arena, M. P., Haggard, K. S. & Yan, X. (2008). Price momentum and idiosyncratic volatility. *Financial Review* 43(2), 159-190.
- Asness, C. S., Moskowitz, T. & Pedersen, L. H. (2013). Value and momentum everywhere. *Journal of Finance* 68(3), 929-985.
- Avramov, D., Chordia, T., Jostova, G. & Philipov, A. (2013). Anomalies and financial distress *Journal of Financial Economics* 108(1), 139-159.
- Bali, T.G., Cakici, N. & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99(2), 426-446.
- Barberis, N., Shleifer, A. & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics* 49(3), 307-343.
- Barroso, P. & Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics* 116(1), 111-120.
- Berk, J. B., Green, R.C. & Naik, V. (1999). Optimal investment, growth options, and security returns. *Journal of Finance* 54(5), 1553-1607.
- Bhootra, A. (2011). Are momentum profits driven by the cross-sectional dispersion in expected stock returns? *Journal of Financial Markets* 14(3), 494-513.'
- Blackburn, D. W. & Cakici, N. (2017). Overreaction and the cross-section of returns: International evidence. *Journal of Empirical Finance* 42 1-14.
- Blitz, D., Huij, J. & Martens, M. (2011). Residual momentum. *Journal of Empirical Finance* 18(3), 506-521.
- Boudoukh, J., Richardson, M. P. & Whitelaw, R. F. (1994). A tale of three schools: Insights on autocorrelations of short-horizon stock returns. *The Review of Financial Studies* 7(3), 539-573.
- Boyer, B., Mitton, T. & Vorkink K. (2010). Expected idiosyncratic skewness. *The Review of Financial Studies* 23(1), 169-202.

- Bulkley, G. & Nawosah, V. (2009). Can the cross-sectional variation in expected stock returns explain momentum? *Journal of Financial and Quantitative Analysis* 44(4), 777-794.
- Chan, L. K. C., Jegadeesh, N. & Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance* 51(5), 1681-1713.
- Chan, W. (1992). A note on time series model specification in the presence of outliers. *Journal of Applied Statistics* 19(1), 117-124.
- Chan, W. (1995). Outliers and financial time series modelling: A cautionary note. *Mathematics and computers in simulation* 39(3-4), 425-.
- Chen, J. & Hong, H. (2002). Discussion of "Momentum and autocorrelation in stock returns". *The Review of Financial Studies* 15(2), 565-574.
- Chordia, T. & Shivakumar, L. (2002). Momentum, business cycle, and Time-varying expected returns. *Journal of Finance* 57(2), 985-1019.
- Chui, A. C. W., Titman, S. & Wei, K. C. J. (2010). Individualism and momentum around the world. *The Journal of Finance* 65(1), 361-392.
- Conrad, J. & Kaul, G. (1998). An anatomy of trading strategies. *The Review of Financial Studies* 11(3), 489-519.
- Conrad, J. & Yavuz, M. D. (2016). Momentum and Reversal: Does What Goes Up Always Come Down? *Review of Finance* 21(2), 555-581.
- Cooper, M. J., Gutierrez Jr., R. C. & Hameed, A. (2004). Market states and momentum. *The Journal of Finance* 59(3), 1345-1365.
- Da, Z., Gurn, U. G. & Warachka, M. (The Review of Financial Studies). Frog in the pan: Continuous information and momentum. *2014* 27(7), 2171-2218.
- Daniel, K., Hirshleifer, D. & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *Journal of Finance* 53(6), 1839-1885.
- Daniel, K. & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics* 122(2), 221-247.
- Daniel, K. & Titman, S. (1999). Market efficiency in an irrational world. *Financial Analysts Journal* 55(6), 28-40.
- Debondt, W. & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance* 40(3), 793.
- Demiguel, V., Nogales, F. J. & Uppal, R. (2014). Stock return serial dependence and out-of-sample portfolio performance. *The Review of Financial Studies* 27(4), 1031-1073.

- Deutsch, S. J., Richards, J. E. & Swain, J. J. (1990). Effects of a single outlier on arma identification. *Communications in Statistics - Theory and Methods* 19(6), 2207-2227.
- Du, D. (2012). Momentum and behavioral finance. *Managerial Finance* 38(4), 364-379.
- Fama, E. F. & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51(1), 55-84.
- Fama, E. F. & French, K. R. (2008). Dissecting anomalies. *Journal of Finance* 63(4), 1653-1678.
- Fama, E. F. & French, K. R. (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies* 29(1), 69-103.
- Frazzini, A. (2006). The disposition effect and underreaction to news. *Journal of Finance* 61(4), 2017-2046.
- Frazzini, A., Israel, R. & Moskowitz, T. J. (2014). Trading costs of asset pricing anomalies. *Fama-Miller Working Paper; Chicago Booth Research Paper No. 14-05*. Retrieved from: <http://ssrn.com/abstract=2294498>
- French, K. R. (2018) U.S. Research Returns Data. Retrieved from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
- Goyal, A. & Jegadeesh, N. (2018). Cross-sectional and time-series tests of return predictability: What is the difference? *Review of Financial Studies* 31(5), 1784-1824.
- Goyal, A. & Wahal, S. (2015). Is momentum an echo? *Journal of Financial and Quantitative Analysis* 50(6), 1237-1267.
- Griffin, J. M., Ji, X. & Martin, J. S. (2005). Global momentum strategies. *The Journal of Portfolio Management* 31(2), 23-39.
- Grinblatt, M. & Han, B. (2005). Prospect theory, mental accounting, and momentum. *Journal of Financial Economics* 78(2), 311-339.
- Grinblatt, M. & Moskowitz, T. J. (2004). Predicting stock price movements from past returns: The role of consistency and tax-loss selling. *Journal of Financial Economics*, 71(3), 541-579.
- Grundy, B. & Martin, J. S. (2001). Understanding the nature of the risks and the source of the rewards to momentum investing. *Review of Financial Studies* 14(1), 29-78.
- Holden, C. & Subrahmanyam, A. (2002). News events, information acquisition, and serial correlation *. *The Journal of Business* 75(1), 1-32.

- Hong, H., Lim, T. & Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance* 55(1), 265-295.
- Hong, H. & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54(6), 2143-2184.
- Hong, H. & Stein, J. C. (2007) Disagreement and the Stock Market. *Journal of Economic Perspectives*, 21(2), 109-128.
- Hou, K., Xue, C. & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies* 28(3), 650-705.
- Hvidkjaer, S. (2006). A trade-based analysis of momentum. *The Review of Financial Studies* 19(2), 457-491.
- Israel, R. & Moskowitz, T. J. (2013). The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics* 108(2), 275-301.
- Jegadeesh, N. & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48(1), 65-91.
- Jegadeesh, N. & Titman, S. (1995). Overreaction, delayed reaction, and contrarian profits. *The review of financial studies* 8(4), 973-993.
- Jegadeesh, N. & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance* 56(2), 699-720.
- Jegadeesh, N. & Titman, S. (2002). Cross-sectional and time-series determinants of momentum returns. *The Review of Financial Studies* 15(1), 143-157.
- Johnson, T. C. (2002). Rational momentum effects. *Journal of Finance* 57(2), 585-608.
- Jostova, G., Nikolova, S., Philipov, A. & Stahel, C. W. (2013). Momentum in corporate bond returns. *The Review of Financial Studies* 26(7), 1649-1693.
- Korajczyk, R. A. & Sadka, R. (2004). Are momentum profits robust to trading costs? *Journal of Finance* 59(3), 1039-1082.
- Lee, C. M. C. & Swaminathan, B. (2000). Price momentum and trading volume. *The Journal of Finance* 55(5), 2017-2069.
- Lehmann, B. (1990). Fads, Martingales, and Market Efficiency. *The Quarterly Journal of Economics*, 105(1), 1-28.
- Lesmond, D. A., Schill, M. J. & Zhou, C. (2004). The illusory nature of momentum profits. *Journal of Financial Economics* 71(2), 349-380.

- Lewellen, J. (2002). Momentum and autocorrelation in stock returns. *The Review of Financial Studies* 15(2), 533-564.
- Lewellen, J. & Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics* 82(2), 289-314.
- Liu, L. & Zhang, L. (2008). Momentum profits, factor pricing, and macroeconomic risk. *Review of Financial Studies* 21(6), 2417-2448.
- Lo, A. W. & MacKinlay, A. G. (1990). When are contrarian profits due to stock market overreaction? *The Review of Financial Studies* 3(2), 175-205.
- Luo, J., Subrahmanyam, A. & Titman, S., (2018). Stock price dynamics with overconfident investors. *Unpublished working paper*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2997001
- McLean, R. D. (2010). Idiosyncratic risk, long-term reversal, and momentum. *Journal of Financial and Quantitative Analysis* 45(4), 883-906.
- McLean, R. D. & Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability? *The Journal of Finance*, 71(1), 5-32.
- Moskowitz, T. J. & Grinblatt, M. (1999). Do industries explain momentum? *Journal of Finance* 54(4), 1249-1290.
- Moskowitz, T. J., Ooi, Y. H. & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics* 104(2), 228-250.
- Newey, W. K & West, K. D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55 (3), 703–708.
- Novy-Marx, R. (2012). Is momentum really momentum? *Journal of Financial Economics* 103(3), 429-453.
- Pan, M., Liano, K. & Huang, G.-C. (2004). Industry momentum strategies and autocorrelations in stock returns. *Journal of Empirical Finance* 11(2), 185-202.
- Pan, M. (2010). Autocorrelation, return horizons, and momentum in stock returns. *Journal of Economics and Finance* 34(3), 284-300.
- Park, K. & Kim, D. (2014). Sources of momentum profits in international stock markets. *Accounting & Finance* 54(2), 567-589.
- Pástor, L. & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of political economy* 111(3), 642-685.
- Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics* 42(1-2), 35-52.

- Rouwenhorst, K. (1998). International momentum strategies. *The Journal of Finance* 53(1), 267-284.
- Ruenzi, S. & Weigert, F. (2018). Momentum and crash sensitivity. *Economics Letters* 165 77.
- Sadka, R. (2006). Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics* 80(2), 309-349.
- Sagi, J. S. & Seasholes, M. S. (2007). Firm-specific attributes and the cross-section of momentum. *Journal of Financial Economics* 84(2), 389-434.
- Shleifer, A. & Vishny, R. (1997). The limits of arbitrage. *Journal of Finance* 52(1), 35-55.
- Sias, R. W. (2007). Reconcilable differences: Momentum trading by institutions. *Financial Review* 42(1), 1-22.
- Stivers, C. & Sun, L. (2010). Cross-sectional return dispersion and time variation in value and momentum premiums. *The Journal of Financial and Quantitative Analysis* 45(4), 987-1014.
- Subrahmanyam, A. (2018). Equity market momentum: A synthesis of the literature and suggestions for future work. *Pacific-Basin Finance Journal* 51 291-296.
- Vayanos, D. & Woolley, P. (2013). An institutional theory of momentum and reversal. *The Review of Financial Studies* 26(5), 1087-1145.
- Welch, I. (2017). Stock return outliers. *Unpublished working paper*. Retrieved 4.3.2019 from <https://ssrn.com/abstract=3068347>.

Appendix 1

STRATEGY AND DECOMPOSITION DERIVATIONS

Decomposition of momentum

Following Lo and MacKinlay (1990):

$$\pi_t = \frac{1}{N} \sum_{i=1}^N (r_{i,t-2}^{11} - \bar{r}_{t-2}^{11}) r_{it}$$

where π_t is the momentum return in month t , N is the number of assets, r_{it} is the return of asset i at time t , $r_{i,t-2}^{11}$ is the cumulative 11-month return of asset i two months prior to t , and \bar{r}_{t-2}^{11} their equally weighted average. This can be rewritten as

$$\begin{aligned} \pi_t &= \frac{1}{N} \sum_{i=1}^N (r_{i,t-2}^{11} r_{it}) - \frac{1}{N} \sum_{i=1}^N (\bar{r}_{t-2}^{11} r_{it}) \\ &= \frac{1}{N} \sum_{i=1}^N (r_{i,t-2}^{11} r_{it}) + \bar{r}_{t-2}^{11} \bar{r}_t \end{aligned}$$

Taking expectations yields

$$\begin{aligned} E[\pi_t] &= \frac{1}{N} \sum_{i=1}^N E[r_{i,t-2}^{11} r_{it}] + E[\bar{r}_{t-2}^{11} \bar{r}_t] \\ &= \frac{1}{N} \sum_{i=1}^N (Cov[r_{i,t-2}^{11}, r_{it}] + \mu_i \mu_i^{11}) - (Cov[\bar{r}_{t-2}^{11}, \bar{r}_t] + \bar{\mu} \bar{\mu}^{11}) \end{aligned}$$

where Cov denotes covariance, μ_i is the unconditional expected monthly return of asset i and μ_i^{11} is the unconditional expected 11-month return of asset i . This can again be rewritten as

$$E[\pi_t] = \frac{1}{N} \text{tr}(\Omega) - \frac{1}{N^2} 1' \Omega_k 1 + \sigma_{\mu\gamma}.$$

Here, 1 denotes an $(N \times 1)$ vector of ones, Ω is the autocovariance matrix $\Omega \equiv E[(r_{t-2}^{11} - \gamma)(r_t - \mu)']$ and the notation $\text{tr}(\cdot)$ denotes the trace of a matrix. In this equation, the first term is the cross-sectional average of the k th-order autocovariances of the individual securities, the second term is the autocovariance of the equal-weighted market index, and the third term $\sigma_{\mu\gamma}$ is the cross-sectional covariance between the unconditional 11-month expected returns, $\gamma = E[r_{i,t}^{11}]$ and the unconditional 1-month mean returns $E[r_{i,t}]$, and represent the contribution of variation in unconditional expected returns. The above equation can be further rearranged into:

$$E[\pi_t] = \frac{N-1}{N^2} \text{tr}(\Omega) - \frac{1}{N^2} [1' \Omega_k 1 - \text{tr}(\Omega)] + \sigma_{\mu\gamma},$$

$$E[\pi_t] = C + O + \sigma_{\mu\gamma}.$$

Derivation of the individual strategy weights

The weights to our strategies are given by

$$w_{i,t} = \frac{1}{N} (r_{t-2}^{11} - \bar{r}_{t-2}^{11}).$$

This can be rewritten as

$$w_{i,t} = \frac{1}{N} \left((r_{t-2}^{11} - \mu_i^{11}) - (\bar{r}_{t-2}^{11} - \bar{\mu}^{11}) \right) + \frac{1}{N} (\mu_i^{11} - \bar{\mu}^{11})$$

$$= w_{o+c} + w_{\sigma},$$

where $w_{o+c} = \frac{1}{N} \left((r_{t-2}^{11} - \mu_i^{11}) - (\bar{r}_{t-2}^{11} - \bar{\mu}^{11}) \right)$, are the weights to a strategy that only exploits time-series dependence in returns of momentum and $w_{\sigma} = \frac{1}{N} (\mu_i^{11} - \bar{\mu}^{11})$ are the weights for a strategy that only depends on the unconditional expected returns of the assets.

The weights for the time-series strategy can further be rewritten as

$$\begin{aligned}
 w_{o+c} &= \frac{1}{N} \left(\frac{N-1}{N} (r_{t-2}^{11} - \mu_i^{11}) - \left(\bar{r}_{t-2}^{11} - \bar{\mu}^{11} - \frac{1}{N} (r_{t-2}^{11} - \mu_i^{11}) \right) \right) \\
 &= \frac{N-1}{N^2} (r_{t-2}^{11} - \mu_i^{11}) - \frac{1}{N} \left(\bar{r}_{t-2}^{11} - \bar{\mu}^{11} - \frac{1}{N} (r_{t-2}^{11} - \mu_i^{11}) \right) \\
 &= w_o + w_c
 \end{aligned}$$

where, $w_o = \frac{N-1}{N^2} (r_{t-2}^{11} - \mu_i^{11})$ are the weights for a strategy that exploits return autocovariances and $w_c = -\frac{1}{N} (\bar{r}_{t-2}^{11} - \bar{\mu}^{11} - \frac{1}{N} (r_{t-2}^{11} - \mu_i^{11}))$ are the weights for a strategy that only depend on cross-covariances among stocks.

Decomposition of the unconditional expected returns strategy

Using the same notation as above,

$$\begin{aligned}
 \sigma_t &= \frac{1}{N} \sum_{i=1}^N (\mu_i^{11} - \bar{\mu}^{11}) r_i \\
 E(\sigma_t) &= \frac{1}{N} \sum_{i=1}^N (E(\mu_i^{11} r_i) - E(\bar{\mu}^{11} r_i)) \\
 &= \frac{1}{N} \sum_{i=1}^N (E(\mu_i^{11} r_i) - \bar{\mu}^{11} \mu_i) \\
 &= cov(\mu_i^{11}, \mu_i)
 \end{aligned}$$

Decomposition of the autocovariance strategy

Using the same notation as above,

$$\begin{aligned}
 O_t &= \frac{1}{N} \sum_{i=1}^N \frac{N-1}{N} (r_{it-2}^{11} - \mu_i^{11}) r_{it} \\
 E(O_t) &= \frac{N-1}{N^2} \sum_{i=1}^N (E(r_{it-2}^{11} r_{it}) - E(\mu_i^{11} r_{it})) \\
 &= \frac{N-1}{N^2} \sum_{i=1}^N (\text{cov}(r_{it-2}^{11} r_{it}) + \mu_i^{11} \mu_i - \mu_i^{11} \mu_i) \\
 &= \frac{N-1}{N^2} \sum_{i=1}^N \text{cov}(r_{it-2}^{11} r_{it}) \\
 &= \frac{N-1}{N^2} \text{tr}(\Omega)
 \end{aligned}$$

Decomposition of the cross-covariance strategy

Using the same notation as above,

$$\begin{aligned}
 C_t &= -\frac{1}{N} \sum_{i=1}^N \left(\bar{r}_{t-2}^{11} - \bar{\mu}^{11} - \frac{1}{N} (r_{it-2}^{11} - \mu_i^{11}) \right) r_{it} \\
 C_t &= -\frac{1}{N} \sum_{i=1}^N (\bar{r}_{t-2}^{11} - \bar{\mu}^{11}) r_{it} + \frac{1}{N^2} \sum_{i=1}^N (r_{it-2}^{11} - \mu_i^{11}) r_{it} \\
 E(C_t) &= -\frac{1}{N} \sum_{i=1}^N (E(\bar{r}_{t-2}^{11} r_{it}) - E(\bar{\mu}^{11} r_{it})) + \frac{1}{N^2} \sum_{i=1}^N (E(r_{it-2}^{11} r_{it}) - E(\mu_i^{11} r_{it}))
 \end{aligned}$$

$$E(C_t) = -\frac{1}{N}(\text{cov}(\bar{r}_{t-2}^{11}\bar{r}_t) + \bar{\mu}^{11}\bar{\mu} - \bar{\mu}^{11}\bar{\mu}) + \frac{1}{N^2}\sum_{i=1}^N(\text{cov}(r_{it-2}^{11}r_{it}))$$

$$E(C_t) = -\frac{\text{cov}(\bar{r}_{t-2}^{11}\bar{r}_t)}{N} + \frac{1}{N^2}\sum_{i=1}^N\left(\text{cov}(r_{it-2}^{11}r_{it})\right)$$

$$E(C_t) = -\frac{1'\Omega 1 - \text{tr}(\Omega)}{N^2}.$$